

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 18:53:39 ; Search time 89.1987 Seconds
(without alignments)
1049.298 Million cell updates/sec

Title: US-10-695-980-32

Perfect score: 1341
Sequence: 1 MSAAALPRADLTATSLIVSG.....LHPTVPMWRLPSTRKGDPA 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1341	100.0	242	2	AAW69535
2	1341	100.0	242	6	ABU97244
3	1332	99.3	242	8	AD061161
4	1332	99.3	242	8	AD061078
5	1332	99.3	242	8	ADQ38247
6	1332	99.3	242	8	ADQ38247
7	1332	99.3	242	8	ADQ38247
8	1332	99.3	242	7	ADQ38247
9	1331	99.3	242	2	AAW69535
10	1331	99.3	242	8	AD061157
11	1331	99.3	242	8	AD061157
12	1331	99.3	242	8	ADQ38243
13	1331	99.3	242	8	ADQ38243
14	1331	99.3	242	8	ADQ38243
15	1336	98.9	299	2	AAW69535
16	1336	98.9	299	5	AAW69535
17	1180	88.0	212	2	AAW69535
18	1180	88.0	212	2	AAW69535
19	1044	77.9	242	2	AAW69535
20	1044	77.9	242	2	AAW69535
21	1044	77.9	242	8	ADQ38245
22	1044	77.9	242	8	ADQ38245
23	1044	77.9	242	8	ADQ38245
24	1044	77.9	242	8	ADQ38245
25	1044	77.9	242	8	ADQ38245

26	1044	77.9	242	8	ADQ38245
27	600	44.7	258	8	ADQ38245
28	600	44.7	258	8	ADQ38245
29	600	44.7	258	8	ADQ38245
30	600	44.7	258	8	ADQ38245
31	600	44.7	258	8	ADQ38245
32	564	42.1	244	6	AAO16024
33	418.5	31.2	262	8	ADQ38245
34	418.5	31.0	262	8	ADQ38245
35	415.5	31.0	262	8	ADQ38245
36	415.5	31.0	262	8	ADQ38245
37	415.5	31.0	262	8	ADQ38245
38	415.5	31.0	262	8	ADQ38245
39	413.5	30.8	262	8	ADQ38245
40	407.5	30.4	258	8	ADQ38245
41	407.5	30.4	258	8	ADQ38245
42	407.5	30.4	258	8	ADQ38245
43	407.5	30.4	258	8	ADQ38245
44	407.5	30.4	258	8	ADQ38245
45	407.5	30.4	258	8	ADQ38245

ALIGNMENTS

RESULT 1
AAW69535
ID AAW69535 standard; protein; 242 AA.

AAW69535;

10-AUG-1999 (first entry)

crw396 gene product beta-carotene beta-oxygenase.

Carotenoid; pigment; canthaxanthin; R1534; crtB; prephytycane synthase;

crtI; phytoene desaturase; crtY; lycopenase cyclase; crtW2396;

beta-carotene beta-oxygenase; food product; fermentation.

Flavobacterium sp.

JP10155497-A.

16-JUN-1998.

02-DEC-1997; 97JP-00348653.

02-DEC-1996; 96EP-00810839.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

WPI; 1998-391048/34.

N-PSDB; AAW40147.

Preparation of carotenoid - comprises fermentation with transformed cell.

Claim 1; Fig 62; 80pp; Japanese.

The invention describes the preparation of carotenoid pigments e.g. canthaxanthin using a cell transformed by a vector having DNA sequences (a) to (e) or substantially homologous sequences. (a) a DNA sequence

(crtB) coding for prephytycane synthase of Flavobacterium sp. R1534;

(c) a DNA sequence (crtI) coding for phytoene desaturase of Flavobacterium

sp. R1534; (d) a DNA sequence (crtY) coding for lycopene cyclase of

Flavobacterium sp. R1534; and (e) a DNA sequence (crtW2396) coding beta-

carotene beta-oxygenase of a microbe E-396 (FERM BP-4283). The carotenoid

or a carotenoid mixture can also be used in preparation of food products.

The method is an improved method of fermentation for carotenoid

production

Sequence 242 AA;

SO

Query Match 100.0%; Score 1341; DB 2; Length 242;
 Best Local Similarity 100.0%; Pred. No. 4.5e-143;
 Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAAHLPKADLTATSLIVSGGIIAAMLALHVALWFLDAAHPILAVANFLGLTWLSVGL 60
 |||||
 DB 1 MSAAHLPKADLTATSLIVSGGIIAAMLALHVALWFLDAAHPILAVANFLGLTWLSVGL 60

QY 61 FIITADHMGSVVPGPRANAMGOLVLMVYAGFSWRKKIVKMAHHRAGTDDDDPDPH 120
 |||||
 DB 61 FIITADHMGSVVPGPRANAMGOLVLMVYAGFSWRKKIVKMAHHRAGTDDDDPDPH 120

QY 121 GGPVRYARFICTGYFGMRGELLPIVITYYALMLGDRMWYVFWPLPSIIASIQLFVFGI 180
 |||||
 DB 121 GGPVRYARFICTGYFGMRGELLPIVITYYALMLGDRMWYVFWPLPSIIASIQLFVFGI 180

QY 181 WLPHRPGHDAFPDRNNAASSRISDPVSLTCTHFGGYHHEHHLHPVPMWRLPSTRTKGD 240
 |||||
 DB 181 WLPHRPGHDAFPDRNNAASSRISDPVSLTCTHFGGYHHEHHLHPVPMWRLPSTRTKGD 240

QY 241 TA 242
 |||||
 DB 241 TA 242

RESULT 2
 ABU97244
 ID ABU97244 standard; protein; 242 AA.
 XX
 AC ABU97244;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Enzyme polypeptide #10.
 XX
 KM Enzyme; hydroxymethylglutaryl coenzyme-A reductase; HMGCoA reductase;
 KM isopentenyl diphosphate isomerase; HMGCoA synthase; mevalonate kinase;
 KM phosphomevalonate kinase; diphosphomevalonate decarboxylase; phyloene;
 KM carotenoid-producing cell; isoprenoid compound; FPP synthase; lycopen;
 KM farneyl diphosphate synthase; carotenoid; beta-carotene; zeaxanthin;
 KM canthaxanthin; astaxanthin; adonixanthin; cryptoxanthin; echinenone;
 KM adonirubin; C-4-isoprenoid compound; nutritional supplement;
 KM food colourant; animal feed pigment.
 XX
 OS Paracoccus carolinifaciens.
 XX
 PN WO200299095-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 05-JUN-2002; 2002WO-EP006171.
 XX
 PR 06-JUN-2001; 2001US-0296299P.
 XX
 PA (HOFF) ROCHE VITAMINS AG.
 XX
 PI Berry A, Bretzel W, Huemelin M, Lopez-Ulibarri R, Mayer AF;
 PI Yeliseev A;
 PI WPI; 2003-156856/15.
 DR N-PSDB; ACA99471.
 XX
 PT New isolated polypeptide used for isoprenoid production, includes amino
 PT acid sequence having e.g. mevalonate kinase, phosphomevalonate kinase or
 PT diphosphomevalonate decarboxylase activity.
 XX
 PS Example 12; Page 295-297; 307pp; English.
 XX
 CC The invention relates to an isolated polypeptide having the activity of
 CC an enzyme such as hydroxymethylglutaryl Coenzyme-A (HMGCoA) reductase,
 CC isopentenyl diphosphate isomerase, HMGCoA synthase, mevalonate kinase,
 CC phosphomevalonate kinase or diphosphomevalonate decarboxylase. The
 CC invention also relates to the production of a carotenoid-producing cell

CC and engineering a bacterium to produce an isoprenoid compound, comprising
 CC culturing a parent bacterium in a medium permitting expression of an
 CC isoprenoid compound and selecting a mutant bacterium from the culture
 CC medium that produces 1.1-1000 times more isoprenoid compound than the
 CC parent bacterium, introducing an expression vector comprising a farneyl
 CC diphosphate (FPP) synthase DNA fragment operably linked to an expression
 CC control sequence into the mutant bacterium, and selecting a bacterium
 CC that contains the expression vector and produces at least 1.1 times more
 CC isoprenoid compound than the selected mutant. The cell is useful for
 CC producing a carotenoid (phyloene, lycopen, beta-carotene, zeaxanthin,
 CC canthaxanthin, astaxanthin, adonixanthin, cryptoxanthin, echinenone or
 CC adonirubin) by culturing the cell under conditions permitting expression
 CC of the polypeptide encoded by the polynucleotide sequence and isolating
 CC the carotenoid from the cell or the medium of the cell. The carotenoids
 CC (C-4- isoprenoid compounds) are used as nutritional supplements,
 CC pharmaceuticals, food colourants and pigments for animal feeds. This
 CC sequence represents a polypeptide of the invention

QY Sequence 242 AA;
 XX
 SQ

Query Match 100.0%; Score 1341; DB 6; Length 242;
 Best Local Similarity 100.0%; Pred. No. 4.5e-143;
 Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAAHLPKADLTATSLIVSGGIIAAMLALHVALWFLDAAHPILAVANFLGLTWLSVGL 60
 |||||
 DB 1 MSAAHLPKADLTATSLIVSGGIIAAMLALHVALWFLDAAHPILAVANFLGLTWLSVGL 60

QY 61 FIITADHMGSVVPGPRANAMGOLVLMVYAGFSWRKKIVKMAHHRAGTDDDDPDPH 120
 |||||
 DB 61 FIITADHMGSVVPGPRANAMGOLVLMVYAGFSWRKKIVKMAHHRAGTDDDDPDPH 120

QY 121 GGPVRYARFICTGYFGMRGELLPIVITYYALMLGDRMWYVFWPLPSIIASIQLFVFGI 180
 |||||
 DB 121 GGPVRYARFICTGYFGMRGELLPIVITYYALMLGDRMWYVFWPLPSIIASIQLFVFGI 180

QY 181 WLPHRPGHDAFPDRNNAASSRISDPVSLTCTHFGGYHHEHHLHPVPMWRLPSTRTKGD 240
 |||||
 DB 181 WLPHRPGHDAFPDRNNAASSRISDPVSLTCTHFGGYHHEHHLHPVPMWRLPSTRTKGD 240

QY 241 TA 242
 |||||
 DB 241 TA 242

RESULT 3
 ADO61161
 ID ADO61161 standard; protein; 242 AA.
 XX
 AC ADO61161;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE P. marcusii ketolase.
 XX
 KM ketocarotenoid; plant; ketolase; petal; flower-specific promoter;
 KM beta-specific promoter; epsilon-cyclase; ec; chromoplast; hydroxylase;
 KM beta-cyclase; bc; astaxanthin; canthaxanthin; echinenone; adonirubin;
 KM adonixanthin; ornamental; food; animal feed; supplement; pigment;
 KM colouring; trout; salmon; shrimp; enzyme.
 XX
 OS Paracoccus marcusii.
 XX
 PN DE10238978-A1.
 XX
 PD 04-MAR-2004.
 XX
 PF 20-AUG-2002; 2002DE-01038978.
 XX
 PR 20-AUG-2002; 2002DE-01038978.
 XX
 PA (SUNG-) SUNGENE GMBH & CO KGAA.
 XX

DR WPI; 2004-215840/21.
 DR N-PSDB; ADO61160.
 PT Method for preparing ketocarotenoids, useful e.g. as food or feed
 PT supplements, by increasing, or introducing, ketolase activity in the
 PT fruits of transgenic plants, also new nucleic acid constructs.
 XX
 PS Disclosure; SEQ ID NO 10; 77bp; German.
 XX
 CC This invention describes a novel method for preparing ketocarotenoids by
 CC culturing genetically modified plants that, in comparison with the wild-
 CC type, have altered ketolase activity in the fruits. The invention also
 CC describes a nucleic acid construct comprising a fruit-specific promoter
 CC linked functionally to a sequence that encodes the altered ketolase,
 CC genetically modified plants in which the fruits have altered ketolase
 CC activity and method for preparing such plants. The plants contain at
 CC least one nucleic acid sequence that encodes the ketolase, especially
 CC under control of a fruit-specific promoter, and particularly they contain
 CC chromoplasts in the fruit. The ketolase-encoding nucleic acid is inserted
 CC by standard methods, then the transgenic plants are cultivated, harvested
 CC and ketocarotenoids are isolated from their fruits. Preferred carotenoids
 CC include, astaxanthin, canthaxanthin, echinenone (or its 3- or 3'-hydroxy
 CC derivatives) adonirubin or adonixanthin. The modified plants with
 CC increased ketolase activity are used as ornamentals as food or animal
 CC feed and for preparation of ketocarotenoids-containing extracts or for
 CC preparing food/feed supplements, e.g., especially where the
 CC ketocarotenoid is astaxanthin, as a pigment for colouring trout, salmon
 CC and shrimps. The transgenic plants produce fruits with increased content
 CC of ketocarotenoids. The invention describes a construct, vector pSKRTO2
 CC which comprises, in pSUN5, a cassette containing the constitutive double
 CC 35S cauliflower mosaic virus promoter; the rbcS chloroplast transit
 CC peptide; the ketolase gene from Haemococcus pluvialis and a terminator.
 CC It was used to transform tomato cells, using Agrobacterium tumefaciens,
 CC and the infected cells regenerated to plants conventionally. One of the
 CC resulting transgenic lines, CS13-24, produced fruits that contained
 CC lycopene, beta-carotene, canthaxanthin, adonirubin or astaxanthin, but
 CC the last three were absent from wild-type fruits (which additionally
 CC contained lutein, not detected in transgenic fruits).
 CC
 XX
 SQ Sequence 242 AA;
 Query Match 99.3%; Score 1332; DB 8; Length 242;
 Best Local Similarity 99.2%; Pred. No. 4.7e-142;
 Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSAAALPKADLTATSLIVSGGIIAAMLAHVHAWFLDAAAPILAVANFLGLTWLSVGL 60
 DB 1 MSAAALPKADLTATSLIVSGGIIAAMLAHVHAWFLDAAAPILAVANFLGLTWLSVGL 60
 QY 61 FIIADHAMGCVVPGRRPRANAAMGQVLMVLYAFSFRKMTVKMAHHRHAGTDDDDPDPH 120
 DB 61 FIIADHAMGCVVPGRRPRANAAMGQVLMVLYAFSFRKMTVKMAHHRHAGTDDDDPDPH 120
 QY 121 GGPVRVYAFIIGTFFCMREGILLPVIVTYVALMGRMMYVWPPLSIASIQLVFPGI 180
 DB 121 GGPVRVYAFIIGTFFCMREGILLPVIVTYVALMGRMMYVWPPLSIASIQLVFPGI 180
 QY 181 WLPHRGHDAPDRNANASSRISDPVSLITCFHFGYVHHHHLLPTVPMWLPESTRTKGD 240
 DB 181 WLPHRGHDAPDRNANASSRISDPVSLITCFHFGYVHHHHLLPTVPMWLPESTRTKGD 240
 QY 241 TA 242
 DB 241 TA 242
 DT 15-JUL-2004 (first entry)

XX
 DB P. marcusii ketolase.
 XX
 KW ketocarotenoid; plant; ketolase; petal; flower-specific promoter;
 KW petal-specific promoter; epsilon-cyclase; ec; chromoplast; hydroxylase;
 KW beta-cyclase; bc; astaxanthin; canthaxanthin; echinenone; adonirubin;
 KW adonixanthin; ornamental; food; animal feed; supplement; pigment;
 KW colouring; trout; salmon; shrimp; enzyme.
 XX
 OS Paracoccus marcusii.
 XX
 PN DE10238980-A1.
 XX
 PD 04-MAR-2004.
 XX
 PF 20-AUG-2002; 2002DE-01038980.
 XX
 PS 20-AUG-2002; 2002DE-01038980.
 XX
 PA (SUNG-) SUNGENE GMBH & CO KGAA.
 XX
 DR WPI; 2004-215842/21.
 DR N-PSDB; ADO61077.
 PT Method for preparing ketocarotenoids, useful e.g. as food or feed
 PT supplements, by increasing, or introducing, ketolase activity in the
 PT petals of transgenic plants, also new nucleic acid constructs.
 XX
 PS Disclosure; SEQ ID NO 10; 140bp; German.
 XX
 CC This invention describes a novel method for preparing ketocarotenoids by
 CC culturing genetically modified plants that, in comparison with the wild-
 CC type, have altered ketolase activity in the petals. The invention also
 CC describes a nucleic acid construct containing a nucleic acid that encodes
 CC ketolase, linked functionally to a flower- or petal-specific promoter,
 CC double-stranded RNA that comprises a sense strand, including a sequence
 CC that is essentially identical with at least part of the RNA epsilon-
 CC cyclase (ec) transcript or the promoter region of the ec gene and an
 CC antisense strand that is essentially complementary to the sense strand,
 CC an transgenic expression cassette comprising a plant-functional promoter
 CC linked to a nucleic acid that transcribes double-stranded RNA,
 CC genetically altered plant in which activity of ketolase in the petals is
 CC increased, if already present in the wild-type or introduced if absent
 CC from the wild type, genetically altered plant that has chromoplasts in
 CC the petals and contains at least one transgenic nucleic acid that encodes
 CC a ketolase. Particularly plants also have increased activity, relative to
 CC wild type, of hydroxylase and/or beta-cyclase (bc) activity, especially
 CC as a result of introducing the appropriate nucleic acid, and plants are
 CC also selected for highest activity of these two enzymes. The transgenic
 CC plants are cultivated, harvested and ketocarotenoids isolated from their
 CC petals. Preferred carotenoids include astaxanthin, canthaxanthin,
 CC echinenone (or its 3- or 3'-hydroxy derivatives), adonirubin or
 CC adonixanthin. The modified plants with increased ketolase activity are
 CC used as ornamentals, as food or animal feed and for preparation of
 CC ketocarotenoids-containing extracts or for preparing food/feed
 CC supplements, e.g., especially where the ketocarotenoid is astaxanthin, as
 CC a pigment for colouring trout, salmon and shrimps. The transgenic plants
 CC have increased content of ketocarotenoids. Vector pSKRTO2 comprises, in
 CC pSUN5, a cassette containing the constitutive double 35S cauliflower
 CC mosaic virus promoter, the rbcS chloroplast transit peptide, the ketolase
 CC gene from Haemococcus pluvialis and a terminator. The vector was used to
 CC transform tomato cells, using Agrobacterium tumefaciens and the infected
 CC cells regenerated to plants conventionally. One of the resulting
 CC transgenic lines, CS13-8, produced orange flowers (yellow in the wild
 CC type) and its petals contained astaxanthin and adonixanthin, both absent
 CC from the wild type.
 XX
 SQ Sequence 242 AA;
 Query Match 99.3%; Score 1332; DB 8; Length 242;
 Best Local Similarity 99.2%; Pred. No. 4.7e-142;
 Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MSAHALPRADLTATSLIVSGGIIAAMLALHVALMFLDAAHPILAVANFLGLTWLSVGL 60
DB 1 MSAHALPRADLTATSLIVSGGIIAAMLALHVALMFLDAAHPILAVANFLGLTWLSVGL 60
QY 61 FIIADNMGSVVPRPRANAMGQLVLMVAGFSMRKVIYKMAHHRAGTDDDDPDFH 120
DB 61 FIIADNMGSVVPRPRANAMGQLVLMVAGFSMRKVIYKMAHHRAGTDDDDPDFH 120
QY 121 GGPVWYARFICTYTGMRGGLLPVIYVYVALMLGDRMNYVFWPLPSILASIQLFVFGI 180
DB 121 GGPVWYARFICTYTGMRGGLLPVIYVYVALMLGDRMNYVFWPLPSILASIQLFVFGI 180
QY 181 WLPHRPGHDAPPDRNANASSRISDPVSLTCTHFGGYHHEHHLHPTVPMWRLPSTRTKGD 240
DB 181 WLPHRPGHDAPPDRNANASSRISDPVSLTCTHFGGYHHEHHLHPTVPMWRLPSTRTKGD 240
QY 241 TA 242
DB 241 TA 242

RESULT 5
ADQ38247
ID ADQ38247 standard; protein; 242 AA.
AC ADQ38247;
XX
XX 23-SEP-2004 (first entry)
DT
XX
XX Paracoccus marcusii ketolase.
DE
XX
XX astaxanthin; plant; animal feed; fodder; fish; crustacean; bird;
XX salmonid; shrimp; crab; hen; duck; geese; flamingo; ketolase;
XX beta-carotene; canthaxanthin; hydroxylase; beta-cyclase; pigmentation;
XX trout; salmon; shrimp.
OS Paracoccus marcusii.
XX
XX DE10258971-A1.
XX
XX 01-JUL-2004.
PD
XX
XX 16-DEC-2002; 2002DE-01058971.
PF
XX 16-DEC-2002; 2002DE-01058971.
PR
XX (SUNG-) SUNGENE GMBH & CO KGAA.
XX
XX WPI; 2004-554601/54.
DR N-PSDB; ADQ38246.
XX
XX Use of astaxanthin-containing plant material, or extracts, from Tagetes
XX for oral administration to animals, particularly for pigmentation of
XX fish, crustacea, birds and their products.
XX
XX Disclosure; SEQ ID NO 10; 145bp; German.
XX
XX This invention describes the novel use of astaxanthin-containing plants
XX (particularly the flower heads or petals) of the genus Tagetes or their
XX parts for oral administration to animals. Methods are also described for
XX a) the preparation of an animal feed composition by mixing standard
XX fodder ingredients with astaxanthin-containing plants, b) for
XX pigmentation of animals, or their products, by oral administration of
XX astaxanthin-containing plants and c) animal feed composition or
XX pigmentation agent that contains astaxanthin-containing plants. The
XX plants, or their parts or extracts, are administered directly to animals,
XX optionally after intermediate processing, or they are formulated with
XX fodder components. Particularly they are administered to fish,
XX crustaceans or birds, specifically salmonids, shrimps, crabs, hens,
XX ducks, geese and flamingoes. The plants of the invention are preferably
XX modified to provide ketolase activity, for conversion of beta-carotene to
XX canthaxanthin, particularly with the highest level of expression in the
XX flowers, e.g. by using a flower-specific promoter. The plants may also

```

CC have increased activity of hydroxylase (for conversion of canthaxanthin
CC to astaxanthin) and/or of beta-cyclase (to increase production of beta-
CC carotene from gamma-carotene). Astaxanthin-containing compositions are
CC used particularly for pigmentation of animals, preferably fish, crustacea
CC and birds, or their products (meat, skin, feathers and eggs), most
CC particularly trout, salmon and shrimp. Genetically modified Tagetes
CC produce larger amounts of astaxanthin than the current source, Adonis
CC aestivalis, and more cheaply. The invention uses expression vector
CC pSKETO2 which includes a cassette consisting of the double 35S promoter,
CC the sequence for the pea rbc transit peptide, the sequence encoding a
CC ketolase (beta-carotene-4-oxygenase) from Haematococcus pluvialis and the
CC polyadenylation signal from cauliflower mosaic virus. It was used to
CC transform leaves of Tagetes, by Agrobacterium-mediated transfer, then
CC these regenerated to plants.

CC Sequence 242 AA;

Query Match 99.3%; Score 1332; DB 8; Length 242;
Best Local Similarity 99.2%; Pred. No. 4,7e-142;
Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MSAHALPRADLTATSLIVSGGIIAAMLALHVALMFLDAAHPILAVANFLGLTWLSVGL 60
DB 1 MSAHALPRADLTATSLIVSGGIIAAMLALHVALMFLDAAHPILAVANFLGLTWLSVGL 60
QY 61 FIIADNMGSVVPRPRANAMGQLVLMVAGFSMRKVIYKMAHHRAGTDDDDPDFH 120
DB 61 FIIADNMGSVVPRPRANAMGQLVLMVAGFSMRKVIYKMAHHRAGTDDDDPDFH 120
QY 121 GGPVWYARFICTYTGMRGGLLPVIYVYVALMLGDRMNYVFWPLPSILASIQLFVFGI 180
DB 121 GGPVWYARFICTYTGMRGGLLPVIYVYVALMLGDRMNYVFWPLPSILASIQLFVFGI 180
QY 181 WLPHRPGHDAPPDRNANASSRISDPVSLTCTHFGGYHHEHHLHPTVPMWRLPSTRTKGD 240
DB 181 WLPHRPGHDAPPDRNANASSRISDPVSLTCTHFGGYHHEHHLHPTVPMWRLPSTRTKGD 240
QY 241 TA 242
DB 241 TA 242

```

```

RESULT 6
ADRO3862
ID ADRO3862 standard; protein; 242 AA.
AC ADRO3862;
XX
XX 21-OCT-2004 (first entry)
DT
XX
XX Paracoccus marcusii ketolase SEQ ID NO: 20.
DE
XX
XX carotenoid; genetically modified; Blakelea; nutrient; pigmentation;
XX antioxidant; transgenic; ketolase; enzyme.
XX
XX Paracoccus marcusii.
XX
XX Synthetic.
XX
XX WO2004063359-A2.
XX
XX 29-JUL-2004.
PD
XX
XX 09-JAN-2004; 2004WO-EP000099.
PF
XX 09-JAN-2003; 2003DE-01000649.
PR 08-SEP-2003; 2003DE-01041271.
XX
XX (BADI ) BASF AG.
XX
XX Matuschek M, Klein D, Heinekamp T, Schmidt A, Brakhage A;
PI Achatz B;
XX
XX WPI; 2004-544088/52.
DR

```


DR N-PSDB; ADR03861.

XX Preparing carotenoids or their precursors useful e.g. in cosmetics,
PT pharmaceuticals, foods and animal feeds, comprises culturing genetically
XX modified *Blakeslea*.

PS Disclosure; SEQ ID NO 20; 486bp; German.

XX The present invention relates to a method of preparing carotenoids or
CC their precursors using genetically modified organisms of the genus
CC *Blakeslea*. The method is used for production of carotenoids, particularly
CC carotenes and xanthophylls, useful as animal and human nutrients, or
CC supplements, cosmetics, dermatological agents and pharmaceuticals,
CC particularly for pigmentation or colouring beverages, but also as
CC antioxidants. The present sequence is a protein used in the
CC exemplification of the invention.

SQ Sequence 242 AA;

Query Match 99.3%; Score 1332; DB 8; Length 242;
Best Local Similarity 99.2%; Pred. No. 4.7e-142;
Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;QY 1 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAWFLDAAHPILAVANFLGLTWLSVGL 60
DB 1 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAWFLDAAHPILAVANFLGLTWLSVGL 60

QY 61 FTIADHMGSVVPGPRRANAMGOLVLMVYGFSGRKMIVKMAHHRHAGTDDDPDFH 120

DB 61 FTIADHMGSVVPGPRRANAMGOLVLMVYGFSGRKMIVKMAHHRHAGTDDDPDFH 120

QY 121 GGPVRMYARFICTYFGMRBGLLPVIVTVYALMLGDRMNVVFWPLPSIIASIQLFVFGI 180

DB 121 GGPVRMYARFICTYFGMRBGLLPVIVTVYALMLGDRMNVVFWPLPSIIASIQLFVFGI 180

QY 181 WLPHRPGHDAFPDRHNAARSSRISDPVSLITCFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240

DB 181 WLPHRPGHDAFPDRHNAARSSRISDPVSLITCFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240

QY 241 TA 242

DB 241 TA 242

RESULT 7

ADR03942

ID ADR03942 standard; protein; 242 AA.

AC ADR03942;

XX 21-OCT-2004 (first entry)

DE P marcusii ketolase.

XX genetically modified; vector; carotenoid production; pigmentation;
KM nutrient; antioxidant; ketolase; enzyme.OS *Paracoccus marcusii*.

PN WO2004063358-A1.

PD 29-JUL-2004.

PF 09-JAN-2004; 2004WO-EP000100.

PR 09-JAN-2003; 2003DE-01000649.

PR 08-SEP-2003; 2003DE-01041272.

(BADT) BASF AG.

XX Matuschek M, Heinekamp T, Schmidt A, Brakhage A;

DR MPI; 2004-544087/52.

DR N-PSDB; ADR03941.

XX Preparing genetically modified *Blakeslea*, useful for preparation of
PT carotenoids, useful as food additives, cosmetics or pharmaceuticals,
XX comprises transformation, optional homokaryotizing, and selection.

PS Disclosure; SEQ ID NO 20; 459bp; German.

XX The present invention relates to a method of preparing a genetically
CC modified organism of the genus *Blakeslea*, which comprises first
CC transforming at least one cell then optionally homokaryotizing the cells
CC so that cells are produced in which the nuclei are all simultaneously
CC altered in one or more genetic characteristics and these alterations are
CC expressed, and finally selection and culture of the modified cell(s). The
CC genetically modified organisms are used for the production of
CC carotenoids, particularly carotenes and xanthophylls, useful as animal
CC and human nutrients, or supplements, cosmetics and pharmaceuticals,
CC particularly for pigmentation or colouring beverages, but also as
CC antioxidants. The present sequence is a protein shown in the
CC exemplification of the invention.

SQ Sequence 242 AA;

Query Match 99.3%; Score 1332; DB 8; Length 242;
Best Local Similarity 99.2%; Pred. No. 4.7e-142;
Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;QY 1 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAWFLDAAHPILAVANFLGLTWLSVGL 60
DB 1 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAWFLDAAHPILAVANFLGLTWLSVGL 60

QY 61 FTIADHMGSVVPGPRRANAMGOLVLMVYGFSGRKMIVKMAHHRHAGTDDDPDFH 120

DB 61 FTIADHMGSVVPGPRRANAMGOLVLMVYGFSGRKMIVKMAHHRHAGTDDDPDFH 120

QY 121 GGPVRMYARFICTYFGMRBGLLPVIVTVYALMLGDRMNVVFWPLPSIIASIQLFVFGI 180

DB 121 GGPVRMYARFICTYFGMRBGLLPVIVTVYALMLGDRMNVVFWPLPSIIASIQLFVFGI 180

QY 181 WLPHRPGHDAFPDRHNAARSSRISDPVSLITCFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240

DB 181 WLPHRPGHDAFPDRHNAARSSRISDPVSLITCFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240

QY 241 TA 242

DB 241 TA 242

RESULT 8

ADQ96834

ID ADQ96834 standard; protein; 648 AA.

AC ADQ96834;

XX 23-SEP-2004 (first entry)

DE CrtWcrty amino acid sequence.

XX carotenoid; transgenic plant; overlapping extension PCR amplification;
KM PCR; crtE; crtB; crtI; crtW; crtZ; astaxanthin; carotene;
XX grain crop; vegetable; crtWcrty.

OS Unidentified.

PN CN1380415-A.

PD 20-NOV-2002.

PF 06-APR-2001; 2001CN-00105878.

PR 06-APR-2001; 2001CN-00105878.

XX

PA (SHAN-) SHANGHAI YONGYE NONGKO BIO ENG CO LTD.

XX Yao Q, Peng R, Xiong A;

XX WPI: 2003-230997/23.

DR N-PSDB; ADQ96833, ADQ96933.

PT Synthesis of related gene for producing carotenoid in transgenic plant,
PT useful for improving quality of grain crops and vegetable.

PS Claim 2; Page 1-2; 36pp; Chinese.

CC The present invention relates to a related gene for producing carotenoid
CC in transgenic plant. The invention utilizes an overlapping extension PCR
CC amplification process and uses six genes of crtB, crtY, crtI, crtW, crtY
CC and crtZ as templates to synthesize the related gene for synthesizing
CC astaxanthin with plant preference code and can make a synthetic gene for
CC expression in a plant to produce the carotene substances of astaxanthin.
CC The synthetic gene can be used to improve the quality of grain crops and
CC vegetables. The present sequence represents a specifically claimed
CC crtWcrtY amino acid sequence, which is used in the exemplification of the
CC present invention.

SQ Sequence 648 AA;

Query Match 99.3%; Score 1332; DB 7; Length 648;

Best Local Similarity 99.2%; Pred. No. 1.7e-141;

Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAFLDAAHPITLAVANFLGLTWLSVGL 60
DB 1 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAFLDAAHPITLAVANFLGLTWLSVGL 60
QY 61 FIHADAMGVSVPGRPRANAMGOLVLMVAGFSWRKMTVKHMAHHRAGTDDPDPDH 120
DB 61 FIHADAMGVSVPGRPRANAMGOLVLMVAGFSWRKMTVKHMAHHRAGTDDPDPDH 120
QY 121 GGPVWYARFIGTYGMRGELLPIVTVVYALMLGDRMYVFWPLPSILASIQLFVFGI 180
DB 121 GGPVWYARFIGTYGMRGELLPIVTVVYALMLGDRMYVFWPLPSILASIQLFVFGI 180
QY 181 WLPHRPGHDAPPDRHNARSSRISDPVSLTCTFHFGYHHEHHLHPTVPMWRLPSTRYGD 240
DB 181 WLPHRPGHDAPPDRHNARSSRISDPVSLTCTFHFGYHHEHHLHPTVPMWRLPSTRYGD 240
QY 241 TA 242
DB 241 TA 242

RESULT 9

AAW82260
ID AAW82260 standard; protein; 242 AA.

XX AAW82260;

DT 17-OCT-2003 (revised)
DT 16-JUL-1999 (first entry)

XX C. utilis crtW protein.

XX HMG-CoA; 3-hydroxy-3-methylglutaryl coenzyme A reductase; crtW;
KW carotenoid.

XX Pichia jadinii.

XX JP10248575-A.

XX 22-SEP-1998.

XX 12-MAR-1997; 97JP-00058012.

XX 12-MAR-1997; 97JP-00058012.

XX (KIRI) KIRIN BREWERY KK.

XX WPI: 1998-560727/48.

DR N-PSDB; AAV73184.

PT Gene useful for increase in carotenoid production - and preparation of
PT carotenoid.

PS Example 2; Fig 20-21; 54pp; Japanese.

CC This invention describes a novel method for the preparation of
CC carotenoids using genes and proteins isolated from *Candida utilis*. The
CC invention specifically describes the isolation of a 3-hydroxy-3-
CC methylglutaryl coenzyme A (HMG-CoA) reductase protein. This sequence
CC represents the *Candida utilis* crtW protein which is used in the method of
CC the invention. (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 242 AA;

Query Match 99.3%; Score 1331; DB 2; Length 242;

Best Local Similarity 98.8%; Pred. No. 6e-142;

Matches 239; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSAHALPRADLTATSLIVSGGIIAAMLAHVHAFLDAAHPITLAVANFLGLTWLSVGL 60
DB 1 MSAHALPRADLTATSLIVSGGIIAAMLAHVHAFLDAAHPITLAVANFLGLTWLSVGL 60
QY 61 FIHADAMGVSVPGRPRANAMGOLVLMVAGFSWRKMTVKHMAHHRAGTDDPDPDH 120
DB 61 FIHADAMGVSVPGRPRANAMGOLVLMVAGFSWRKMTVKHMAHHRAGTDDPDPDH 120
QY 121 GGPVWYARFIGTYGMRGELLPIVTVVYALMLGDRMYVFWPLPSILASIQLFVFGI 180
DB 121 GGPVWYARFIGTYGMRGELLPIVTVVYALMLGDRMYVFWPLPSILASIQLFVFGI 180
QY 181 WLPHRPGHDAPPDRHNARSSRISDPVSLTCTFHFGYHHEHHLHPTVPMWRLPSTRYGD 240
DB 181 WLPHRPGHDAPPDRHNARSSRISDPVSLTCTFHFGYHHEHHLHPTVPMWRLPSTRYGD 240
QY 241 TA 242
DB 241 TA 242

RESULT 10

ADO61157
ID ADO61157 standard; protein; 242 AA.

XX ADO61157;

DT 15-JUL-2004 (first entry)

XX A. aurantiacum ketolase.

XX ketocarotenoid; plant; ketolase; petal; flower-specific promoter;
KW petal-specific promoter; epsilon-cyclase; ec; chromoplas; hydroxylase;
KW beta-cyclase; bc; astaxanthin; canthaxanthin; echinone; adonirubin;
KW adonixanthin; ornamental; food; animal feed; supplement; pigment;
KW colouring; trout; salmon; shrimp; enzyme.

XX Agrobacterium aurantiacum.

XX DE10238978-A1.

XX 04-MAR-2004.

XX 20-AUG-2002; 2002DE-01038978.

XX 20-AUG-2002; 2002DE-01038978.

XX 20-AUG-2002; 2002DE-01038978.

XX (SUNG-) SUNGENE GMBH & CO KGAA.

DR WPI; 2004-215840/21.
 N-PSDB; ADO61156.
 XX
 PT Method for preparing ketocarotenoids, useful e.g. as food or feed
 PT supplements, by increasing, or introducing, ketolase activity in the
 PT fruits of transgenic plants, also new nucleic acid constructs.
 XX
 PS Disclosure; SEQ ID NO 6; 77bp; German.
 XX
 CC This invention describes a novel method for preparing ketocarotenoid by
 CC culturing genetically modified plants that, in comparison with the wild-
 CC type, have altered ketolase activity in the fruits. The invention also
 CC describes a nucleic acid construct comprising a fruit-specific promoter
 CC linked functionally to a sequence that encodes the altered ketolase,
 CC genetically modified plants in which the fruits have altered ketolase
 CC activity and method for preparing such plants. The plants contain at
 CC least one nucleic acid sequence that encodes the ketolase, especially
 CC under control of a fruit-specific promoter, and particularly they contain
 CC chromoplasts in the fruit. The ketolase-encoding nucleic acid is inserted
 CC by standard methods, then the transgenic plants are cultivated, harvested
 CC and ketocarotenoids are isolated from their fruits. Preferred carotenoids
 CC include astaxanthin, canthaxanthin, echinone (or its 3- or 3'-hydroxy
 CC derivatives) adonirubin or adonixanthin. The modified plants with
 CC increased ketolase activity are used as ornamentals as food or animal
 CC feed and for preparation of ketocarotenoid-containing extracts or for
 CC preparing food/feed supplements, e.g., especially where the
 CC ketocarotenoid is astaxanthin, as a pigment for colouring trout, salmon
 CC and shrimps. The transgenic plants produce fruits with increased content
 CC of ketocarotenoids. The invention describes a construct, vector p53KET02
 CC which comprises, in pSUN5, a cassette containing the constitutive double
 CC 35S cauliflower mosaic virus promoter; the rbcS chloroplast transit
 CC peptide; the ketolase gene from *Haemococcus pluvialis* and a terminator.
 CC It was used to transform tomato cells, using *Agrobacterium tumefaciens*,
 CC and the infected cells regenerated to plants conventionally. One of the
 CC resulting transgenic lines, CS13-24, produced fruits that contained
 CC lycopene, beta-carotene, canthaxanthin, adonirubin or astaxanthin, but
 CC the last three were absent from wild-type fruits (which additionally
 CC contained lutein, not detected in transgenic fruits).
 CC
 XX
 SO Sequence 242 AA;
 Query Match 99.3%; Score 1331; DB 8; Length 242;
 Best Local Similarity 98.8%; Pred. No. 6e-142;
 Matches 239; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSAAHPKADLTSLVSGGIIAAMTALVHALVFLDAAHPILAVANLGLTWTLSVGL 60
 DB 1 MSAAHPKADLTSLVSGGIIAAMTALVHALVFLDAAHPILAVANLGLTWTLSVGL 60
 QY 61 FIIAHDMGSSVVGPRPRAAANGOLVLMYAGPSMRKMTVKMAHHRRHAGTDDDDPFDH 120
 DB 61 FIIAHDMGSSVVGPRPRAAANGOLVLMYAGPSMRKMTVKMAHHRRHAGTDDDDPFDH 120
 QY 121 GGFVRYARITGYTFGRREGILLPVITYVALMGDRMWTVPWPLPSILASIQLEVFGI 180
 DB 121 GGFVRYARITGYTFGRREGILLPVITYVALMGDRMWTVPWPLPSILASIQLEVFGI 180
 QY 181 WLFHRGHDAPPRRNARRSRIDPVSLLTCFHFPGYHNEHLLPVPWMLPSTRTKGD 240
 DB 181 WLFHRGHDAPPRRNARRSRIDPVSLLTCFHFPGYHNEHLLPVPWMLPSTRTKGD 240
 QY 241 TA 242
 DB 241 TA 242

RESULT 11

ADO61074
 ID ADO61074 standard; protein; 242 AA.
 AC ADO61074;
 XX
 DT 15-JUL-2004 (first entry)

XX
 DE A. aurantiacum ketolase.
 XX
 KW ketocarotenoid; plant; ketolase; petal; flower-specific promoter;
 KW petal-specific promoter; epsilon-cyclase; ec; chromoplast; hydroxylase;
 KW beta-cyclase; bc; astaxanthin; canthaxanthin; echinone; adonirubin;
 KW adonixanthin; ornamental; food; animal feed; supplement; pigment;
 KW colouring; trout; salmon; shrimp; enzyme.
 XX
 OS *Agrobacterium aurantiacum*.
 XX
 PN DE10238980-A1.
 XX
 PD 04-MAR-2004.
 XX
 PF 20-AUG-2002; 2002DE-01038980.
 PR 20-AUG-2002; 2002DE-01038980.
 PR (SUNG-) SUNGENE GMEH & CO KGMA.
 PA
 DR WPI; 2004-215842/21.
 DR N-PSDB; ADO61073.
 XX
 PT Method for preparing ketocarotenoids, useful e.g. as food or feed
 PT supplements, by increasing, or introducing, ketolase activity in the
 PT petals of transgenic plants, also new nucleic acid constructs.
 XX
 PS Disclosure; SEQ ID NO 6; 140bp; German.
 XX
 CC This invention describes a novel method for preparing ketocarotenoid by
 CC culturing genetically modified plants that, in comparison with the wild-
 CC type, have altered ketolase activity in the petals. The invention also
 CC describes a nucleic acid construct containing a nucleic acid that encodes
 CC ketolase, linked functionally to a flower- or petal-specific promoter,
 CC double-stranded RNA that comprises a sense strand, including a sequence
 CC that is essentially identical with at least part of the RNA epsilon-
 CC cyclase (ec) transcript or the promoter region of the ec gene and an
 CC antisense strand that is essentially complementary to the sense strand,
 CC an transgenic expression cassette comprising a plant-functional promoter
 CC linked to a nucleic acid that transcribes double-stranded RNA,
 CC genetically altered plant in which activity of ketolase in the petals is
 CC increased, if already present in the wild-type or introduced if absent
 CC from the wild type, genetically altered plant that has chromoplasts in
 CC the petals and contains at least one transgenic nucleic acid that encodes
 CC a ketolase. Particularly plants also have increased activity, relative to
 CC wild type, of hydroxylase and/or beta-cyclase (bc) activity, especially
 CC as a result of introducing the appropriate nucleic acid, and plants are
 CC also selected for highest activity of these two enzymes. The transgenic
 CC plants are cultivated, harvested and ketocarotenoids isolated from their
 CC petals. Preferred carotenoids include astaxanthin, canthaxanthin,
 CC echinone (or its 3- or 3'-hydroxy derivatives), adonirubin or
 CC adonixanthin. The modified plants with increased ketolase activity are
 CC used as ornamentals, as food or animal feed and for preparation of
 CC ketocarotenoid-containing extracts or for preparing food/feed
 CC supplements, e.g., especially where the ketocarotenoid is astaxanthin, as
 CC a pigment for colouring trout, salmon and shrimps. The transgenic plants
 CC have increased content of ketocarotenoids. Vector p53KET02 comprises, in
 CC pSUN5, a cassette containing the constitutive double 35S cauliflower
 CC mosaic virus promoter, the rbcS chloroplast transit peptide, the ketolase
 CC gene from *Haemococcus pluvialis* and a terminator. The vector was used to
 CC transform tomato cells, using *Agrobacterium tumefaciens* and the infected
 CC cells regenerated to plants conventionally. One of the resulting
 CC transgenic lines, CS13-8, produced orange flowers (yellow in the wild
 CC type) and its petals contained astaxanthin and adonixanthin, both absent
 CC from the wild type.
 CC
 XX
 SO Sequence 242 AA;
 Query Match 99.3%; Score 1331; DB 8; Length 242;
 Best Local Similarity 98.8%; Pred. No. 6e-142;
 Matches 239; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MSAHLPRKADLVAATSLIVSGGIIAAMLAHVLAHWLDDAAAHPIIAVANDLGLTWLSVGL	60
Db	1	MSAHLPRKADLVAATSLIVSGGIIAAMLAHVLAHWLDDAAAHPIIAVANDLGLTWLSVGL	60
Qy	61	FTIADAHMGSVVPGHPRANAAMGQLVLMVYAGFSNRKMI VKMAIHHRAAGTDDDPDFDH	120
Db	61	FTIADAHMGSVVPGHPRANAAMGQLVLMVYAGFSNRKMI VKMAIHHRAAGTDDDPDFDH	120
Qy	121	GGPVHWYARFICTYEGMRBGLLPVIVTYYALMLGPRMYYVWPPLPSITASTQLPVFGI	180
Db	121	GGPVHWYARFICTYEGMRBGLLPVIVTYYALMLGPRMYYVWPPLPSITASTQLPVFGI	180
Qy	181	WLPHRPGHDAPFDRHNAASRSRISDPVSLITCFHFGGYHNEHHLAPTVPMWRLPSTRTKGD	240
Db	181	WLPHRPGHDAPFDRHNAASRSRISDPVSLITCFHFGGYHNEHHLAPTVPMWRLPSTRTKGD	240
Qy	241	TA 242	
Db	241	TA 242	

RESULT 12
ADQ38243
ID ADQ38243 standard; protein; 242 AA

AC	ADQ30243;
XX	
DT	23-SEP-2004 (first entry)

DE *Agrobacterium aurantiacum* ketolase.

KM agsaxanthin; plant; animal feed; fodder; fish; crustacean; bird;
 KM salmonid; shrimp; crab; hen; duck; geese; flamingo; ketolase;
 KM beta-carotene; canthaxanthin; hydroxylase; beta-cyclase; pigmenta-
 KM tion; trout; salmon; shrimp.

OS *Agrobacterium aurantiacum*.

PN DE10258971-A1.

PD 01-JUL-2004.

PF 16-DEC-2002; 2002DE-01058971.

PR 16-DEC-2002; 2002DE-01058971.

PA (SUNG-) SUNGENE GMBH & CO KGAA.

DR WPI; 2004-554601/54.

XX

PT use of astaxanthin-containing plant material, or extracts, from algae or other sources for oral administration to animals, particularly for pigmentation of fish, crustacea, birds and their products.

PS Disclosure; SEQ ID NO 6; 145pp; German

This invention describes the novel use of astaxanthin-containing plants (particularly the flower heads or petals) of the genus *Tagetes* or their parts for oral administration to animals. Methods are also described for a) the preparation of an animal feed composition by mixing standard fodder ingredients with astaxanthin-containing plants, b) for pigmentation of animals, or their products, by oral administration of astaxanthin-containing plants and c) animal feed composition or pigmentation agent that contains astaxanthin-containing plants. The plants, or their parts or extracts, are administered directly to animals, optionally after intermediate processing, or they are formulated with fodder components. Particularly they are administered to fish, crustaceans or birds, specifically salmonids, shrimps, crabs, hens, ducks, geese and flamingoes. The plants of the invention are preferably modified to provide ketolase activity, for conversion of beta-carotene to canthaxanthin, particularly with the highest level of expression in the flowers, e.g. by using a flower-specific promoter. The plants may also

CC have increased activity of hydroxylase (for conversion of canthaxanthin
CC to astaxanthin) and/or of beta-cyclase (to increase production of beta-
CC carotene from gamma-carotene). Astaxanthin-containing compositions are
CC used particularly for pigmentation of animals, preferably fish, crustacea
CC and birds, or their products (meat, skin, feathers and eggs), most
CC particularly trout, salmon and shrimp. Genetically modified *Tageetes*
CC produce larger amounts of astaxanthin than the current source, *Adonis*
CC aestivalls, and more cheaply. The invention uses expression vector
CC pSNET2 which includes a cassette consisting of the double 35S promoter,
CC the sequence for the pea rbc transit peptide, the sequence encoding a
CC ketolase (beta-carotene-4-oxygenase) from *Haematococcus pluvialis* and the
CC polyadenylation signal from cauliflower mosaic virus. It was used to
CC transform leaves of *Tageetes*, by *Agrobacterium*-mediated transfer, then
CC these regenerated to plants.

Query Match	99.3%	Score 1331;	DB 8;	Length 242;
Best Local Similarity	98.8%	Pred. No. 6e-142;		
Matches 239; Conservative	2;	Mismatches 1;	Indels 0;	Gaps 0

QY 1 MSAHALPKADLTATSLIVSGIIAAMLALHVALWFLDAAHPILAVANFLGLTWSVGL 60

Db 1 MSAHALPKADLTATSLIVSGGIIAAWLALHVHALWFLDAAHPILAIANFLGLTWLSVGL 60

QY 61 FIAHDAMHGSVPPGRPRANAMGOLVWL YAGFSWRKMI VIKHMAHHRHAGTDDDDPFDH 12

Db 61 FIAHADMGSVVPGRPANAMGQLVWL YAGFSWRKMI VKMAHHRHAGTDDDDPDEDH 12

QY 121 GGPVRYARFIGTYFGWREGLLPVI VT VYALMLGDRMNVVFWPLPSIIASIQLFEVGI 18

Db 121 GGPVRYARFIGTYFGWREGLLPVIVTYVALILGDRWMYVFWPLPSILASIQLFVEGT 18

QY 181 WLPHPGDAFPDRHNARSSRISDPVSLTCTFHFGGYHHEHLHPTVPWWRLPSTRTKGD 24

Db 181 WLPHRPGHDAFPDRHNARSSRISDPVSLTCTFHFGGYHHEHLHPVPWWRLPSTRTKGD 24

QY 241 TA 242

Db 241 TA 242

1111

ADR03858

XX

XX

XX

100

KW antioxidant; transgenic; ketolase; enzyme.

05 *Agrobacterium aurantiacum*.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

3 XX
3
4
3
4

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PR 08-SEP-2003; 2003DE-01041271.

PA (BADI) BASF AG.

PI Matuschek M, Klein D, Heinekamp T, Schmidt A, Brakhage A,

XX 2004 E44000/EC

DR N-PSDB; ADR03857.
XX
PT Preparing carotenoids or their precursors useful e.g. in cosmetics,
PT pharmaceuticals, foods and animal feeds, comprises culturing genetically
XX modified *Blakeslea*.
XX
PS Disclosure; SEQ ID NO 16; 486bp; German.
XX
CC The present invention relates to a method of preparing carotenoids or
CC their precursors using genetically modified organisms of the genus
CC *Blakeslea*. The method is used for production of carotenoids, particularly
CC carotenes and xanthophylls, useful as animal and human nutrients, or
CC supplements, cosmetics, dermatological agents and pharmaceuticals,
CC particularly for pigmentation or colouring beverages, but also as
CC antioxidants. The present sequence is a protein used in the
CC exemplification of the invention.
XX
SQ Sequence 242 AA;
XX
Query Match 99.3%; Score 1331; DB 8; Length 242;
Best Local Similarity 98.8%; Pred. No. 6e-142;
Matches 239; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MSNHALPKADLTATSLIVSGGIIAAMWLAHVHAWLWFLDAAAHPIIAVANFLGLTWLSVGL 60
Db 1 MSNHALPKADLTATSLIVSGGIIAAMWLAHVHAWLWFLDAAAHPIIAVANFLGLTWLSVGL 60
XX
QY 61 FIIAHDMAGSVVPGPRRANAMGOLVLMYAGFSWRKMIYKMAHHRHAGTDDDDPDFH 120
Db 61 FIIAHDMAGSVVPGPRRANAMGOLVLMYAGFSWRKMIYKMAHHRHAGTDDDDPDFH 120
XX
QY 121 GGPVRYARFIQTYFGMRGGLLPVIVTYVALMDGRWYVVFMPPLPSIIASIQLFVFGI 180
Db 121 GGPVRYARFIQTYFGMRGGLLPVIVTYVALMDGRWYVVFMPPLPSIIASIQLFVFGI 180
XX
QY 181 WLPHRGHDAPFDRNNAARSSRISDPVSLTCTFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240
Db 181 WLPHRGHDAPFDRNNAARSSRISDPVSLTCTFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240
XX
QY 241 TA 242
Db 241 TA 242
XX
RESULT 14
ADR03938
ID ADR03938 standard; protein; 242 AA.
XX
AC ADR03938;
XX
DT 21-OCT-2004 (first entry)
XX
DE A aurantiacum ketolase.
XX
KM genetically modified; vector; carotenoid production; pigmentation;
KM nutrient; antioxidant; ketolase; enzyme.
XX
OS *Agrobacterium aurantiacum*.
OS Synthetic.
XX
XX WO2004063358-A1.
XX
XX 29-JUL-2004.
XX
XX 09-JAN-2004; 2004WO-EP000100.
XX
XX 09-JAN-2003; 2003DE-01000649.
PR 08-SEP-2003; 2003DE-01041272.
XX
XX (BAD1) BASF AG.
XX
XX Matuechek M, Heinekamp T, Schmidt A, Brakhage A;
XX

DR WPI; 2004-544087/52.
DR N-PSDB; ADR03937.
XX
XX
PT Preparing genetically modified *Blakeslea*, useful for preparation of
PT carotenoids, useful as food additives, cosmetics or pharmaceuticals,
PT comprises transformation, optional homokaryotizing, and selection.
XX
XX
PS Disclosure; SEQ ID NO 16; 459bp; German.
XX
XX
CC The present invention relates to a method of preparing a genetically
CC modified organism of the genus *Blakeslea*, which comprises first
CC transforming at least one cell then optionally homokaryotizing the cells
CC so that cells are produced in which the nuclei are all simultaneously
CC altered in one or more genetic characteristics and these alterations are
CC expressed, and finally selection and culture of the modified cell(s). The
CC genetically modified organisms are used for the production of
CC carotenoids, particularly carotenes and xanthophylls, useful as animal
CC and human nutrients, or supplements, cosmetics and pharmaceuticals,
CC particularly for pigmentation or colouring beverages, but also as
CC antioxidants. The present sequence is a protein shown in the
CC exemplification of the invention.
XX
SQ Sequence 242 AA;
XX
Query Match 99.3%; Score 1331; DB 8; Length 242;
Best Local Similarity 98.8%; Pred. No. 6e-142;
Matches 239; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MSNHALPKADLTATSLIVSGGIIAAMWLAHVHAWLWFLDAAAHPIIAVANFLGLTWLSVGL 60
Db 1 MSNHALPKADLTATSLIVSGGIIAAMWLAHVHAWLWFLDAAAHPIIAVANFLGLTWLSVGL 60
XX
QY 61 FIIAHDMAGSVVPGPRRANAMGOLVLMYAGFSWRKMIYKMAHHRHAGTDDDDPDFH 120
Db 61 FIIAHDMAGSVVPGPRRANAMGOLVLMYAGFSWRKMIYKMAHHRHAGTDDDDPDFH 120
XX
QY 121 GGPVRYARFIQTYFGMRGGLLPVIVTYVALMDGRWYVVFMPPLPSIIASIQLFVFGI 180
Db 121 GGPVRYARFIQTYFGMRGGLLPVIVTYVALMDGRWYVVFMPPLPSIIASIQLFVFGI 180
XX
QY 181 WLPHRGHDAPFDRNNAARSSRISDPVSLTCTFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240
Db 181 WLPHRGHDAPFDRNNAARSSRISDPVSLTCTFHFGYHHEHHLAPTVPMWRLLPSTRTKGD 240
XX
QY 241 TA 242
Db 241 TA 242
XX
RESULT 15
AAW98198
ID AAW98198 standard; protein; 299 AA.
XX
AC AAW98198;
XX
DT 05-JUL-1999 (first entry)
XX
DE SSU/beta-carotene ketolase fusion.
XX
XX Btea-carotene ketolase; crtW gene; carotenoid; xanthophyll; pigment;
KM astaxanthin; zeaxanthin; alpha-carotene; beta-carotene; phytoene;
KM echinenone; canthaxanthin; transgenic plant; oilseed plant; seed oil;
XX A. aurantiacum.
XX
XX Unidentified.
XX
XX WO9907867-A1.
XX
XX 18-FEB-1999.
XX
XX 06-AUG-1998; 98WO-US016466.
XX
XX 08-AUG-1997; 97US-00908758.
XX

```
XX (CALJ ) CALGENE LLC.
PA
XX
XX Shewmaker CK;
PI
XX
XX WPI, 1999-180495/15.
DR
XX N-PSDB; AAX25068.
XX
PT Altering xanthophyll content of seeds by transformation - used to produce
XX seed oils of increased carotenoid content, e.g. Brassica and cotton.
XX
PS Example 1; Fig 16; 92pp; English.
XX
XX The present sequence is a fusion between a 5SU leader and a beta-carotene
CC ketolase encoded by the crtW gene (see AAX25068) of A. aurantiacum. The
CC xanthophyll content of seeds can be altered by transforming cells of a
CC host plant with at least one construct containing: (i) a transcription
CC initiation region of a gene that is preferentially expressed in seeds;
CC (ii) the sequence for a plastid transit peptide; (iii) DNA from a
CC carotenoid synthesis gene coding region, and (iv) a transcription
CC terminator. The transformed cells are regenerated to plants and these, or
CC their progeny, grown to produce seeds. The method is especially used to
CC increase the carotenoid content in oilseed plants. Coexpression of beta-
CC carotene hydroxylase (see AAX98197) and beta-carotene ketolase results in
CC increased astaxanthin and zeaxanthin content in the seeds of transformed
CC plants. Coexpression of beta-carotene ketolase with phytoene synthase
CC (see AAX25063) results in increased levels of alpha-carotene, beta-
CC carotene and phytoene, as well as echinenone and canthaxanthin
XX
SQ Sequence 299 AA;
Query Match 98.9%; Score 1326; DB 2; Length 299;
Best Local Similarity 98.3%; Pred. No. 2.9e-141;
Matches 238; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAFLDAAHPITLAVANFLGLTWLSVGL 60
DB 58 MSAHALPKADLTATSLIVSGGIIAAMLAHVHAFLDAAHPITLAVANFLGLTWLSVGL 117
QY 61 FTIAHDAMHGSVYVGRPRANAAMGOLVWLYAGFSMRKMIIVKMAHHRHAGTDDPDPDH 120
DB 118 FTIAHDAMHGSVYVGRPRANAAMGOLVWLYAGFSMRKMIIVKMAHHRHAGTDDPDPDH 177
QY 121 GGPVRYARFI GTYFGMRBGLLPVIVTYAALMLGDRMYVVFVWPLPSILASIQLFVFGI 180
DB 178 GGPVRYARFI GTYFGMRBGLLPVIVTYAALMLGDRMYVVFVWPLPSILASIQLFVFGI 237
QY 181 WLPHRRGHDAPFDRNNAARSSRISDPVSLTCTHFGGYHHEHHILPTVPMWRLLPSTRTKGD 240
DB 238 WLPHRRGHDAPFDRNNAARSSRISDPVSLTCTHFGGYHHEHHILPTVPMWRLLPSTRTKGD 297
QY 241 TA 242
DB 298 TA 299
```

Search completed: June 17, 2005, 19:33:05
Job time : 90.1987 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 19:21:35 ; Search time 23.6571 Seconds
(without alignments)
763.623 Million cell updates/sec

Title: US-10-695-980-32

Perfect score: 1341

Sequence: 1 MSAHALPKADLTATSLIVSG.....LHPTVPMWRLPSTRTKGDRA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*\n2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*\n3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*\n4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*\n5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*\n6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1341	100.0	242	3	US-08-980-832-32
2	1341	100.0	242	2	US-09-920-923B-32
3	1183	88.2	212	2	US-08-663-310-2
4	1183	88.2	212	2	US-09-006-491-2
5	1183	88.2	212	2	US-09-335-919-2
6	1044	77.9	242	2	US-08-663-310-9
7	1044	77.9	242	2	US-09-006-491-9
8	1044	77.9	242	2	US-09-335-919-9
9	1044	77.9	242	3	US-08-980-832-29
10	1044	77.9	242	4	US-09-920-923B-29
11	354.5	26.4	329	2	US-08-562-535C-4
12	354.5	26.4	329	2	US-08-742-605D-4
13	354.5	26.4	329	3	US-09-259-294-4
14	342.5	25.5	288	2	US-08-632-434-8
15	342.5	25.5	313	2	US-08-632-434-6
16	342.5	25.5	320	2	US-08-632-434-2
17	342.5	25.5	320	2	US-08-632-434-4
18	180.5	13.5	279	4	US-09-902-540-16352
19	180	13.4	129	4	US-09-902-540-13791
20	142.5	10.6	307	4	US-09-902-540-10475
21	135.5	10.1	453	3	US-08-244-205-13
22	135.5	10.1	453	2	US-09-161-994A-10
23	135.5	10.1	453	5	PCT-US92-10284-13
24	134	10.0	355	4	US-09-252-991A-26758
25	128	9.5	256	4	US-09-902-540-10535
26	127.5	9.5	335	4	US-09-328-352-5349
27	127.5	9.5	435	5	PCT-US94-01321-12

28	127	9.5	386	2	US-08-244-205-2	Sequence 2, Appl
29	127	9.5	386	5	US-09-161-994A-11	Sequence 11, Appl
30	127	9.5	386	3	PCT-US92-10284-2	Sequence 2, Appl
31	125.5	9.4	329	3	US-09-161-994A-6	Sequence 6, Appl
32	125.5	9.4	378	2	US-08-244-205-7	Sequence 7, Appl
33	125.5	9.4	378	5	PCT-US92-10284-7	Sequence 7, Appl
34	125.5	9.4	404	2	US-08-244-205-9	Sequence 9, Appl
35	125.5	9.4	404	5	PCT-US92-10284-9	Sequence 9, Appl
36	125.5	9.4	421	4	US-09-857-583B-4	Sequence 4, Appl
37	125	9.3	377	3	US-09-161-994A-12	Sequence 12, Appl
38	125	9.3	460	3	US-09-161-994A-7	Sequence 7, Appl
39	120.5	9.0	436	3	US-09-161-994A-9	Sequence 9, Appl
40	120.5	9.0	446	2	US-08-244-205-5	Sequence 5, Appl
41	120.5	9.0	446	3	US-09-161-994A-5	Sequence 5, Appl
42	120.5	9.0	446	5	PCT-US92-10284-5	Sequence 5, Appl
43	120.5	9.0	446	5	PCT-US94-01321-10	Sequence 10, Appl
44	118.5	8.8	380	2	US-08-244-205-11	Sequence 11, Appl
45	118.5	8.8	380	3	US-09-161-994A-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT 1
US-08-980-832-32
; Sequence 32, Application US/08980832B
; Patent No. 6291204
; GENERAL INFORMATION:
; APPLICANT: Tasankees, Luis
; APPLICANT: Teygankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/08/980,832B
; CURRENT FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 32
; LENGTH: 242
; TYPE: PRT
; ORGANISM: E-396
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Unknown
US-08-980-832-32

Query Match      100.0%; Score 1341; DB 3; Length 242;
Best Local Similarity 100.0%; Pred. No. 2.2e-140; Indels 0; Gaps 0;
Matches 242; Conservative 0; Mismatches 0;

QY      1 MSAHALPKADLTATSLIVSGIIAAMLAHVAMFLDAAAPILAVANFLGITWLSVGL 60
        |||||||
Db      1 MSAHALPKADLTATSLIVSGIIAAMLAHVAMFLDAAAPILAVANFLGITWLSVGL 60

QY      61 FTIADAMGVSVPGRPRANAMGQVLMVLVAFGSRKXIVGMAHRRAGTDDPDPDH 120
        |||||||
Db      61 FTIADAMGVSVPGRPRANAMGQVLMVLVAFGSRKXIVGMAHRRAGTDDPDPDH 120

QY      121 GGPVRYAFITGTGMRGGLPVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 180
        |||||||
Db      121 GGPVRYAFITGTGMRGGLPVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 180

QY      181 WLPHRGHAFFDRNRASRISDPVSLTCHFGYHHEHLHPVPMWRLPSTRTKGD 240
        |||||||
Db      181 WLPHRGHAFFDRNRASRISDPVSLTCHFGYHHEHLHPVPMWRLPSTRTKGD 240

QY      241 TA 242
        ||
Db      241 TA 242

RESULT 2
US-09-920-923B-32
; Sequence 32, Application US/09920923B
; Patent No. 6677134
```

GENERAL INFORMATION:
APPLICANT: Pasamontes, Luis
APPLICANT: Tsyanakov, Yuri
TITLE OF INVENTION: Fermentative Carotenoid Production
FILE REFERENCE: 15464 US (C38435/125944)
CURRENT APPLICATION NUMBER: US/09/920,923B
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 08/980,832
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32
LENGTH: 242
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: E-396
US-09-920-923B-32

Query Match 100.0%; Score 1341; DB 4; Length 242;
Best Local Similarity 100.0%; Pred. No. 2,2e-140;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAHALPKADLTATSLIVSGGIIAAMTALHVAHAFLLDAAPILAVANFLGLTWSGL 60
DB 1 MSAHALPKADLTATSLIVSGGIIAAMTALHVAHAFLLDAAPILAVANFLGLTWSGL 60
QY 61 FLIAHDAMHGSVPPGRPRANAMGQVLMLYAGFSRKMIVKMAHHRAGTDDDPDH 120
DB 61 FLIAHDAMHGSVPPGRPRANAMGQVLMLYAGFSRKMIVKMAHHRAGTDDDPDH 120
QY 121 GSPVWYAFITGYFGMRGLLPVIVTYAALMDGRMYVFWPLPSILASIQLFVFGI 180
DB 121 GSPVWYAFITGYFGMRGLLPVIVTYAALMDGRMYVFWPLPSILASIQLFVFGI 180
QY 181 WLPHPGHDAFPDRNNAASSRISDPVSLTCTHFGGYHHEHLHPVPMWRLPSTRTKGD 240
DB 181 WLPHPGHDAFPDRNNAASSRISDPVSLTCTHFGGYHHEHLHPVPMWRLPSTRTKGD 240
QY 241 TA 242
DB 241 TA 242

RESULT 3

US-08-663-310-2
Sequence 2, Application US/08663310

GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5811273ihiko
APPLICANT: Kondo, Keiji
APPLICANT: KAJIWARA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-663-310-2

Query Match 88.2%; Score 1183; DB 2; Length 212;
Best Local Similarity 98.6%; Pred. No. 5,8e-123;
Matches 209; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 31 VHALFLLDAAPILAVANFLGLTWSGLFLIAHDAMHGSVPPGRPRANAMGQVLML 90
DB 1 VHALFLLDAAPILAVANFLGLTWSGLFLIAHDAMHGSVPPGRPRANAMGQVLML 60
QY 91 YAGFSRKMIVKMAHHRAGTDDDPDHGSPVWYAFITGYFGMRGLLPVIVTY 150
DB 61 YAGFSRKMIVKMAHHRAGTDDDPDHGSPVWYAFITGYFGMRGLLPVIVTY 120
QY 151 ALMLGDRMYVFWPLPSILASIQLFVGIWLPHPGHDAFPDRNNAASSRISDPVSLT 210
DB 121 ALMLGDRMYVFWPLPSILASIQLFVGIWLPHPGHDAFPDRNNAASSRISDPVSLT 180
QY 211 CFHFGYHHEHLHPTVPMWRLPSTRTKGDTA 242
DB 181 CFHFGYHHEHLHPTVPMWRLPSTRTKGDTA 212

RESULT 4

US-09-006-491-2
Sequence 2, Application US/09006491

GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5972690ihiko
APPLICANT: Kondo, Keiji
APPLICANT: KAJIWARA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,491

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-09-006-491-2

Query Match 88.2%; Score 1183; DB 2; Length 212;
Best Local Similarity 98.6%; Pred. No. 5,8e-123;
Matches 209; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 31 VHALWFLDAAHPILAVANFLGLTWLSVGLFIADHMGSVVPGPRANAMGQLVLM 90
1 VHALWFLDAAHPILAVANFLGLTWLSVGLFIADHMGSVVPGPRANAMGQLVLM 60
DB 91 YAGFSWRKRIIVKMAHHRAGTDDPDPHGSPVRYAFIGTYFGMRGGLLPVIVTY 150
61 YAGFSWRKRIIVKMAHHRAGTDDPDPHGSPVRYAFIGTYFGMRGGLLPVIVTY 120
QY 151 AALMGDRMYVFWPLPSILASIQLFVFGIWLPHRPGHDAFPDRNARSRIISDPVSLT 210
121 AALMGDRMYVFWPLPSILASIQLFVFGIWLPHRPGHDAFPDRNARSRIISDPVSLT 180
DB 211 CFHFGYHHEHHLPTVPMWRLPSTRTKGD 242
181 CFHFGYHHEHHLPTVPMWRLPSTRTKGD 212

RESULT 5
US-09-335-919-2
Sequence 2, Application US/09335919
Patent No. 6150130
GENERAL INFORMATION:
APPLICANT: MISA, No. 6150130hiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIYAMA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
TITLE OF INVENTION: XANTHOPHYLLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,919
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-09-335-919-2

Query Match 88.2%; Score 1183; DB 3; Length 212;
Best Local Similarity 98.6%; Pred. No. 5,8e-123;
Matches 209; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 31 VHALWFLDAAHPILAVANFLGLTWLSVGLFIADHMGSVVPGPRANAMGQLVLM 90
1 VHALWFLDAAHPILAVANFLGLTWLSVGLFIADHMGSVVPGPRANAMGQLVLM 60
DB 91 YAGFSWRKRIIVKMAHHRAGTDDPDPHGSPVRYAFIGTYFGMRGGLLPVIVTY 150
61 YAGFSWRKRIIVKMAHHRAGTDDPDPHGSPVRYAFIGTYFGMRGGLLPVIVTY 120
QY 151 AALMGDRMYVFWPLPSILASIQLFVFGIWLPHRPGHDAFPDRNARSRIISDPVSLT 210
121 AALMGDRMYVFWPLPSILASIQLFVFGIWLPHRPGHDAFPDRNARSRIISDPVSLT 180
DB 211 CFHFGYHHEHHLPTVPMWRLPSTRTKGD 242
181 CFHFGYHHEHHLPTVPMWRLPSTRTKGD 212

RESULT 6
US-08-663-310-9
Sequence 9, Application US/08663310
Patent No. 5811273
GENERAL INFORMATION:
APPLICANT: MISA, No. 5811273hiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIYAMA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
TITLE OF INVENTION: XANTHOPHYLLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA


```

CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,919
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 242 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-335-919-9

Query Match      77.9%; Score 1044; DB 3; Length 242;
Best Local Similarity 77.5%; Pred. No. 1.7e-107;
Matches 179; Conservative 17; Mismatches 35; Indels 0; Gaps 0;

QY 12 TATSLIVSGGIIAAMTALHVAHLMFLDAAAHPIIAVANFLGLTWLSVGLPIIAHDMHGS 71
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 12 TIVNIGLTAAILLCWLVLAFTMLDAAAHPLAVLCLAGLTWLSVGLPIIAHDMHGS 71
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 72 VVPGRRANAAAGOLVLMYAGFSWPKLAKMTHRHAGTDNDPFGGCVRYWYARFI 131
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 72 VVPGRRANAAAGOLVLMYAGFSWPKLAKMTHRHAGTDNDPFGGCVRYWYARFI 131
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 132 GTYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 132 GTYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 132 STYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 132 STYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 192 PRRHNAARSRIIDPVSLTCTFHFGYHHEHHLHPVPMWRLPSTRKTGDTA 242
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 192 PRRHNAARSRIIDPVSLTCTFHFGYHHEHHLHPVPMWRLPSTRKTGDTA 242
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 9
US-08-980-832-29
Sequence 29, Application US/08980832B
Patent No. 6291204
GENERAL INFORMATION:
APPLICANT: Pasamontes, Luis
APPLICANT: Teyganikov, Yuri
TITLE OF INVENTION: Improved Fermentative Carotenoid Production
FILE REFERENCE: Improved Fermentative Carotenoid
```

```

CURRENT APPLICATION NUMBER: US/08/980,832B
CURRENT FILING DATE: 1997-12-01
NUMBER OF SEQ ID NOS: 66
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 29
LENGTH: 242
TYPE: PRT
ORGANISM: Alcaligenes PC-1
US-08-980-832-29

Query Match      77.9%; Score 1044; DB 3; Length 242;
Best Local Similarity 77.5%; Pred. No. 1.7e-107;
Matches 179; Conservative 17; Mismatches 35; Indels 0; Gaps 0;

QY 12 TATSLIVSGGIIAAMTALHVAHLMFLDAAAHPIIAVANFLGLTWLSVGLPIIAHDMHGS 71
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 12 TIVNIGLTAAILLCWLVLAFTMLDAAAHPLAVLCLAGLTWLSVGLPIIAHDMHGS 71
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 72 VVPGRRANAAAGOLVLMYAGFSWPKLAKMTHRHAGTDNDPFGGCVRYWYARFI 131
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 72 VVPGRRANAAAGOLVLMYAGFSWPKLAKMTHRHAGTDNDPFGGCVRYWYARFI 131
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 132 GTYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 132 GTYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 132 STYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 132 STYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 192 PRRHNAARSRIIDPVSLTCTFHFGYHHEHHLHPVPMWRLPSTRKTGDTA 242
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 192 PRRHNAARSRIIDPVSLTCTFHFGYHHEHHLHPVPMWRLPSTRKTGDTA 242
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 10
US-09-920-923B-29
Sequence 29, Application US/09920923B
Patent No. 6677134
GENERAL INFORMATION:
APPLICANT: Pasamontes, Luis
APPLICANT: Teyganikov, Yuri
TITLE OF INVENTION: Fermentative Carotenoid Production
FILE REFERENCE: 15464 US (C38435/125944)
CURRENT APPLICATION NUMBER: US/09/920,923B
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 08/980,832
PRIOR FILING DATE: 1997-12-01
NUMBER OF SEQ ID NOS: 66
SOFTWARE: Patentin version 3.1
SEQ ID NO 29
LENGTH: 242
TYPE: PRT
ORGANISM: Alcaligenes PC-1
US-09-920-923B-29

Query Match      77.9%; Score 1044; DB 4; Length 242;
Best Local Similarity 77.5%; Pred. No. 1.7e-107;
Matches 179; Conservative 17; Mismatches 35; Indels 0; Gaps 0;

QY 12 TATSLIVSGGIIAAMTALHVAHLMFLDAAAHPIIAVANFLGLTWLSVGLPIIAHDMHGS 71
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 12 TIVNIGLTAAILLCWLVLAFTMLDAAAHPLAVLCLAGLTWLSVGLPIIAHDMHGS 71
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 72 VVPGRRANAAAGOLVLMYAGFSWPKLAKMTHRHAGTDNDPFGGCVRYWYARFI 131
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 72 VVPGRRANAAAGOLVLMYAGFSWPKLAKMTHRHAGTDNDPFGGCVRYWYARFI 131
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 132 GTYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 132 GTYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 132 STYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 132 STYFGMRBGLLPVITVYVAAALMDGRWYVFWPPLSIASIQLFVFGIWLPHRPGHDAF 191
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 192 PRRHNAARSRIIDPVSLTCTFHFGYHHEHHLHPVPMWRLPSTRKTGDTA 242
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 192 PRRHNAARSRIIDPVSLTCTFHFGYHHEHHLHPVPMWRLPSTRKTGDTA 242
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
```

[illegible]

RESULT 12
US-08-742-605D-4
; Sequence 4, Application US/08742605D
; Patent No. 5965795
; GENERAL INFORMATION:

1 APPLICANT: Joseph Hirschberg, Tamar Lotan and
 2 APPLICANT: Mark Harker
 3 TITLE OF INVENTION: Polynucleotide molecule from
 4 TITLE OF INVENTION: Hematococcus pluvialis encoding a
 5 TITLE OF INVENTION: polypeptide having a beta-C-4-oxygenase
 6 TITLE OF INVENTION: activity for biotechnological production of
 7 TITLE OF INVENTION: (3S,3'S) astaxanthin.
 8 NUMBER OF SEQUENCES: 4
 9 CORRESPONDENCE ADDRESS:
 10 ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein
 11 STREET: 2940 Birchtree space lane
 12 CITY: Silver Spring
 13 STATE: Maryland
 14 COUNTRY: United States of America
 15 ZIP: 20906
 16 COMPUTER READABLE FORM:
 17 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
 18 COMPUTER: Chicom NB5500/386SX
 19 OPERATING SYSTEM: MS DOS version 6.2,
 20 OPERATING SYSTEM: Windows version 3.11
 21 SOFTWARE: word for windows version 2.0,
 22 SOFTWARE: converted to ASCII
 23 CURRENT APPLICATION DATA:
 24 APPLICATION NUMBER: US/08/742,605D
 25 FILING DATE: Oct. 28, 1996
 26 CLASSIFICATION: 435
 27 PRIOR APPLICATION DATA:
 28 APPLICATION NUMBER: 08/562,535
 29 FILING DATE: No. 5965795, 24, 1995
 30 ATTORNEY/AGENT INFORMATION:
 31 NAME: Friedman, Mark M.
 32 REGISTRATION NUMBER: 33,883
 33 REFERENCE/DOCKET NUMBER: 325/12
 34 TELECOMMUNICATION INFORMATION:
 35 TELEPHONE: 972-3-5625553
 36 TELEFAX: 972-3-5625554
 37 TELEX:
 38 INFORMATION FOR SEQ ID NO: 4:
 39 SEQUENCE CHARACTERISTICS:
 40 LENGTH: 329 amino acids
 41 TYPE: amino acid
 42 TOPOLOGY: linear
 43 US-08-742-605D-4
 44 Query Match 26 4%; Score 354.5; DB 2; Length 329;
 45 Best Local Similarity 31.5%; Pred. NO. 6.3e-31;
 46 Matches 84; Conservative 44; Mismatches 104; Indels 35; Gaps 7

RESULT 13
US-09-259-294-4
; Sequence 4, Application US/09259294
; Patent No. 6218599

```

1 GENERAL INFORMATION:
2 APPLICANT: Joseph Hirschberg, Tamar Lotan and
3 APPLICANT: Mark Harker
4 TITLE OF INVENTION: Polynucleotide molecule from
5 TITLE OF INVENTION: Haematococcus pluvialis encoding a
6 TITLE OF INVENTION: polypeptide having a beta-C-4-oxygenase
7 TITLE OF INVENTION: activity for biotechnological production of
8 TITLE OF INVENTION: (3S,3'S) astaxanthin.
9 NUMBER OF SEQUENCES: 4
10 CORRESPONDENCE ADDRESS:
11 ADDRESSER: Mark M. Friedman c/o Anthony Castorina
12 STREET: 2001 Jefferson Davis Highway, Suite 207
13 CITY: Arlington
14 STATE: Virginia
15 COUNTRY: United States of America
16 ZIP: 22202
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
20 COMPUTER: Chicony NB5500/386SX
21 OPERATING SYSTEM: MS DOS version 6.2,
22 OPERATING SYSTEM: Windows version 3.11
23 SOFTWARE: word for Windows version 2.0,
24 SOFTWARE: converted to ASCII
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: US/09/259,294
27 FILING DATE:
28
29 CLASSIFICATION:
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: 08/742,605
32 FILING DATE: Oct. 28, 1996
33 APPLICATION NUMBER: 08/562,535
34 FILING DATE: No. 6218599, 24, 1995
35 ATTORNEY/AGENT INFORMATION:
36 NAME: Friedman, Mark M.
37 REGISTRATION NUMBER: 33, 883
38 REFERENCE/DOCKET NUMBER: 325/12
39 TELECOMMUNICATION INFORMATION:
40 TELEPHONE: 972-3-562553
41 TELEFAX: 972-3-562554
42 TELEX:
43 INFORMATION FOR SEQ ID NO: 4:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 329 amino acids
46 TYPE: amino acid
47 TOPOLOGY: linear
48
49 US-09-259-294-4
50
51 Query Match 26.4%; Score 354.5; DB 3; Length 329;
52 Basic Local Similarity 31.5%; Pred. No. 6.3e-31;
53 Matches 84; Conservative 44; Mismatches 104; Indels 35; Gaps 7
54
55 QY 2 SAHALPKKADLTPTSIVSGGIITAAIMTLAHVLMFL-----DAAPHILA 46
56 DB 56 NAKKRPBDDTKGTYMAIR--VIGSAAVATGIAIPIQIKLPTSIDQLHMLPVSATATQVSG 113
57 QY 47 VAN-----FLGTLMSVGLFIIAHDHMGSVVGRPRANAAMQQLVLMYAGFSWRKM 99
58 DB 114 TSSLDIVVFLVLEFLYTGLFITTHDAMHGIIAARNRQNLDFLRGVCSISYAFWDMYL 173
59 QY 100 IVKKAHNRHAG-TDDDPEDHGGP--VRMYRFLGTGFMREGILLPIYIVYVYALMGD 156
60 DB 174 HRKHENHNHGEVKKDDPFHRGNIGIYVWFAFSFMSYSNMWQFARLMTWTVWQDLGAP 233
61 QY 157 RMYVVFVWLPRLPIIASIOFLVGVILPHRP-----GHDAPRDHNARSSRISDPVSL 209
62 DB 234 MANLVFAAARILTAAPRLFTGTITMFKRFEAGASGSPAVMMWMSKRTSQASDLVFL 293
63 QY 210 TCFHFGYHNEHILHPTVPMWLPSTR 236
64 DB 294 TCYHF-DLHWEHNRWPAFWMELPNCR 319

```

```

US-08-632-434-8
; Sequence 8, Application US/08632434
; Patent No. 5910433
; GENERAL INFORMATION:
; APPLICANT: KAJIWARA, Susumu
; APPLICANT: MISAWA, No. 5910433hhko
; APPLICANT: KONDO, Keiji
; TITLE OF INVENTION: KETO GROUP-INTRODUCING ENZYME, DNA
; TITLE OF INVENTION: CODING THEREFOR AND METHOD FOR PRODUCING KETOCAROTENOIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/632,434
; FILING DATE: 23-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 95/047266
; FILING DATE: 07-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 94/223798
; FILING DATE: 19-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 94/198775
; FILING DATE: 23-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 81356/106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-632-434-8

Query Match      25.5%; Score 342.5; DB 2; Length 288;
Best Local Similarity 31.7%; Pred. No. 1,1e-29;
Matches 84; Conservative 39; Mismatches 107; Indels 35; Gaps 8;

3 AAALPKADLTATSLIVSGIIAAWLAHAL-----W-----FLDA 40
          :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db AYKPPASDAQKITWALP--IIIGTWAVFLHAIQIRLPTSMQLHWLPVSEATQLGG 73
OY 41 AHPLIAVANPLGLTWSVGFLITAHDMHGSSVYGRPRRAAMQVLNLYAGSWEKMI 100
          ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 74 SSLIHIAVFVLEFLVTGTFITTHDHMGITALRRHLNDLLGNITISLAYWFYSMLH 133
OY 101 VKHMAHRHAG-TDDDPDPHGGP--VRMYARFIGTYFGWRREGILLPVITVYALMDGR 157
          |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 134 RKHEHHNHNGEVEKDDDFHKNGDLVPWPASPSTMSLMQPRLKMWAVVMQMLGAPM 193
OY 158 WMYVFWPPLPSIIASIQLFVFGIMLPHP-----GHDAEPDRNRNASRSIDFVSLITC 211
          :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 194 ANLLVFMAAAPLIISAFRLFPGGYLPXKPEKPAAGSQWA-WFRAKTSIASDVMSFLTCC 252
OY 212 FHFGCTHHEHHLHPTVWWRLPSTR 236
          :|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Db 253 YHF-DLHWEHHRMPAPMWQLPHCR 276

RESULT 15

US-08-632-434-6

/ Sequence 6, Application US/08632434

/ Patent No. 5910433

/ GENERAL INFORMATION:

/ APPLICANT: KAJIMAWA, Susumu

/ APPLICANT: MISAWA, No. 5910433;hiko

/ APPLICANT: KONDO, Keiji

/ TITLE OF INVENTION: KETO GROUP-INTRODUCING ENZYME, DNA

/ TITLE OF INVENTION: CODING THEREFOR AND METHOD FOR PRODUCING KETOCAROTENOID

/ NUMBER OF SEQUENCES: 11

/ CORRESPONDENCE ADDRESSES:

/ ADDRESSEE: Foley & Lardner

/ STREET: 3000 K Street, N.W., Suite 500

/ CITY: Washington

/ STATE: D.C.

/ COUNTRY: USA

/ ZIP: 20007-5109

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent in Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/632,434

/ FILING DATE: 23-APR-1996

/ CLASSIFICATION: 435

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: JP 95/047266

/ FILING DATE: 07-MAR-1995

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: JP 94/223798

/ FILING DATE: 19-SEP-1994

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: JP 94/198775

/ FILING DATE: 23-AUG-1994

/ ATTORNEY/AGENT INFORMATION:

/ NAME: BENT, Stephen A.

/ REGISTRATION NUMBER: 29,768

/ REFERENCE/DOCKET NUMBER: 81356/106

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (202) 672-5300

/ TELEFAX: (202) 672-5399

/ TELETYPE: 904136

/ INFORMATION FOR SEQ ID NO: 6:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 313 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

/ US-08-632-434-6

Query Match 25.5%; Score 342.5; DB 2; Length 313;

Best Local Similarity 31.7%; Pred. No. 1.3e-29;

Matches 84; Conservative 39; Mismatches 107; Indels 35; Gaps 8;

QY 3 AHAALPKADLTATSLVSGGIAMALAHYAL-----W-----FLDAA 40

Db 41 AYKPPASDAKGTMTALT--IIGTWTAVFLHAIFOIRLPTSMQDHLVPSBATQAQLGGS 98

QY 41 AHPILAVANFLGLTWISVGLFITIAHDMGSVVGRPRANAAGOLVWLVAQFSWRKMI 100

Db 99 SSLIHIAAVFIYIEFLYTGIFITTHDMGTTALRRQLNDLIGNICISLIAMFDYSMLH 158

QY 101 VKMAHHRHAG--TDDDPDPDHGSP--VRWYARFITYFGWRBGLPLVIVYVYALMGDR 157

Db 159 RKHWEHHRHNTGEVKDPDFHKNGPLVPPFASFMSTYSLMQPARLAMWAVVWMLGAPM 218

QY 158 WMTVVVFWPLPSIIASIQLVFGIWLPHRP-----GHDAFPDRHNAARSRISDPVSLTLC 211

Db 219 ANLVFMAAAILSAFLFYFGTYLPHKPEEGPAAGSQVMA-WFPAKTSSEASDVMSFLTC 277

QY 212 FAFGGYHHEHHLAPTVPMWRLPSTR 236

Db 278 YHF-DLHWEHHRMPAPMWQLPHCR 301

Search completed: June 17, 2005, 19:39:11
Job time : 24.6571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 19:37:17 ; Search time 82.6058 Seconds
(without alignments)
1124.898 Million cell updates/sec

Title: US-10-695-980-32

Perfect score: 1341

Sequence: 1 MSAHALPKADLTATSLIVSG.....LHPTVPMRLPSTRKGDYA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues 1714042

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppa/US05_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubppa/US11_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1341	100.0	242	US-09-920-923-32	Sequence 32, Appl
2	1341	100.0	242	US-10-166-225A-181	Sequence 181, Appl
3	1341	100.0	242	US-10-695-980-32	Sequence 32, Appl
4	1341	99.3	242	US-10-997-844-33	Sequence 33, Appl
5	1044	77.9	242	US-09-920-923-29	Sequence 29, Appl
6	1044	77.9	242	US-10-695-980-29	Sequence 29, Appl
7	564	42.1	244	US-10-466-656-39	Sequence 39, Appl
8	564	42.1	244	US-10-848-307-39	Sequence 39, Appl
9	354.5	26.4	329	US-09-791-687A-4	Sequence 4, Appl
10	146	10.9	453	US-10-310-154-466	Sequence 466, Appl
11	146	10.9	453	US-10-732-923-359	Sequence 359, Appl

12	146	10.9	454	US-10-732-923-4850	Sequence 4850, Ap
13	145	10.8	363	US-10-732-923-4730	Sequence 4730, Ap
14	144	10.7	268	US-10-369-493-15558	Sequence 15558, A
15	144	10.7	280	US-10-369-493-15931	Sequence 15931, A
16	144	10.7	280	US-10-369-493-16296	Sequence 16296, A
17	144	10.7	298	US-10-200-545-90	Sequence 90, Appl
18	143.5	10.7	493	US-10-732-923-4742	Sequence 4742, Ap
19	141	10.5	447	US-10-425-115-310937	Sequence 310937, A
20	139.5	10.4	398	US-10-732-923-4809	Sequence 4809, Ap
21	139	10.4	341	US-10-314-657-20	Sequence 20, Appl
22	139	10.4	341	US-10-473-193-20	Sequence 20, Appl
23	137	10.2	347	US-10-369-493-20100	Sequence 20100, A
24	135.5	10.1	350	US-10-369-493-19077	Sequence 19077, A
25	135.5	10.1	366	US-10-115-571A-28	Sequence 28, Appl
26	135.5	10.1	453	US-10-732-923-4852	Sequence 4852, Ap
27	135	10.1	424	US-10-732-923-4892	Sequence 4892, Ap
28	135	10.1	526	US-10-732-923-4853	Sequence 4853, Ap
29	134.5	10.0	385	US-10-732-923-4795	Sequence 4795, Ap
30	134	10.0	337	US-10-369-493-19438	Sequence 19438, A
31	134	10.0	349	US-10-732-923-4891	Sequence 4891, A
32	134	10.0	354	US-10-115-571A-29	Sequence 29, Appl
33	134	10.0	443	US-10-425-115-310903	Sequence 310903, A
34	134	10.0	443	US-10-732-923-4810	Sequence 4810, Ap
35	134	10.0	443	US-10-732-923-4811	Sequence 4811, Ap
36	131.5	9.8	373	US-10-115-571A-27	Sequence 27, Appl
37	131.5	9.8	398	US-10-732-923-4807	Sequence 4807, Ap
38	131.5	9.8	398	US-10-732-923-4808	Sequence 4808, Ap
39	130	9.7	340	US-10-369-493-11971	Sequence 11971, A
40	130	9.7	340	US-10-732-923-4774	Sequence 4774, Ap
41	129	9.6	292	US-10-732-923-4816	Sequence 4816, Ap
42	128.5	9.6	349	US-10-115-571A-33	Sequence 33, Appl
43	128.5	9.6	383	US-10-732-923-4804	Sequence 4804, Ap
44	128.5	9.6	383	US-10-732-923-4805	Sequence 4805, Ap
45	128.5	9.6	383	US-10-732-923-4806	Sequence 4806, Ap

ALIGNMENTS

RESULT 1
US-09-920-923-32
; Sequence 32, Application US/09920923
; Publication No. US20030022273A1
; GENERAL INFORMATION:
; APPLICANT: Teyganov, Luis
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/09/920,923
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 242
; TYPE: PRT
; ORGANISM: E-396
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Unknown
US-09-920-923-32

QY	1	MSAHALPKADLTATSLIVSGGIIAAMLAHVLAFLDAAAPILAVANFLGLTWLSVGL	60
DB	1	MSAHALPKADLTATSLIVSGGIIAAMLAHVLAFLDAAAPILAVANFLGLTWLSVGL	60
QY	61	FIHAADMGSVVPGPRPRANAMGOLVLMVLYGFSMRKMTVGHMAHRRAGTDDPDFFH	120
DB	61	FIHAADMGSVVPGPRPRANAMGOLVLMVLYGFSMRKMTVGHMAHRRAGTDDPDFFH	120

Qy	121	GGGVNRWARRIGFVFGMRBGLLLPVLYVYALMGDRMMVVFVMPBLSIIASIQLFVFGI	180
	121	GGGVNRWARRIGFVFGMRBGLLLPVLYVYALMGDRMMVVFVMPBLSIIASIQLFVFGI	180
Db	121	GGGVNRWARRIGFVFGMRBGLLLPVLYVYALMGDRMMVVFVMPBLSIIASIQLFVFGI	180
Qy	181	WLPHRPGCHDAFPDRNHAASSRISDPVSLITCFHFGYGHENHLHPYVWNRRLPSTRTKGD	240
	181	WLPHRPGCHDAFPDRNHAASSRISDPVSLITCFHFGYGHENHLHPYVWNRRLPSTRTKGD	240
Qy	241	TA	242
	241	TA	242
Db	241	TA	242

RESULT 2
US-10-166-225A-181
Sequence 181, Application US/10166225A
Publication No. US20030148416A1
GENERAL INFORMATION:
APPLICANT: BERRY, Alan
APPLICANT: BRETZEL, Werner
APPLICANT: HUMBELIN, Martinus
APPLICANT: LOPEZ-OLIBARRI, Rual
APPLICANT: MAYER, Anne F.
APPLICANT: YEISEEV, Alexei A.
TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
FILE REFERENCE: C38435/121966
CURRENT FILING DATE: 2002-06-05
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn version 3.1
SEQ ID NO 181
LENGTH: 242
TYPE: PR1
ORGANISM: Paracoccus carotinifaciens E-396
US-10-166-225A-181

Query Match	100.0%	Score 1341;	DB 14;	Length 242;
Best Local Similarity	100.0%	Pred. No. 4,4e-128;		
Matches	242;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0
Qy	1	MSAHALPRADLTATISLIYSGGIIAAMLAHLVHALFLLDAAHPITAAVNFGLITLSTGL	60	
Db	1	MSAHALPRADLTATISLIYSGGIIAAMLAHLVHALFLLDAAHPITAAVNFGLITLSTGL	60	
Qy	61	FLIADAHNGSVVPRPRANAAMGQVLMLYAGFSMRKATVGHMAHNRAGTDDDPDDFH	120	
Db	61	FLIADAHNGSVVPRPRANAAMGQVLMLYAGFSMRKATVGHMAHNRAGTDDDPDDFH	120	
Qy	121	GGPVRWYARFICTYFGMRBGLLPVITVYVALMLGDRMVTVFWPLPSLIASIQLFVFGI	180	
Db	121	GGPVRWYARFICTYFGMRBGLLPVITVYVALMLGDRMVTVFWPLPSLIASIQLFVFGI	180	
Qy	181	WLPNHPGDAFPDRNNAASSRISDVSLLTCFPGYHNHNHLPTVPWMLPSTRTKGD	240	
Db	181	WLPNHPGDAFPDRNNAASSRISDVSLLTCFPGYHNHNHLPTVPWMLPSTRTKGD	240	
Qy	241	TA 242		
Db	241	TA 242		

RESULT 3
US-10-695-980-32
Sequence 32, Application US/10695980
Publication No US200400505410A1
GENERAL INFORMATION:
APPLICANT: Pasmontes, Luis
APPLICANT: Teygankov, Yuri
TITLE OF INVENTION: Fermentative Carotenoid Production
FILE REFERENCE: 15464 US (C98435/125944)
CURRENT APPLICATION NUMBER: US/10/695,980
CURRENT FILING DATE: 2003-10-29

```

? PRIOR APPLICATION NUMBER: US/09/920,923
? PRIOR FILING DATE: 2001-08-02
? PRIOR APPLICATION NUMBER: 08/980,832
? PRIOR FILING DATE: 1997-12-01
? NUMBER OF SEQ ID NOS: 66
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 32
? LENGTH: 242
? TYPE: PRT
? ORGANISM: Unknown
? FEATURE:
? OTHER INFORMATION: E-396
US-10-695-980-32

```

	Query Match	100.0%;	Score 1341;	DB 15;	Length 242;
	Best Local Similarity	100.0%;	Pred. No. 4,4e-128;		
Matches	242;	Conservative	0;	Mismatches	0;
				Indels	0;
QY	1	MSAHLPRADLTATSLIVSGGIIAAMLAHYHALMFLDAAHPILAVANPFGLTWLSGL	60		
DB	1	MSAHLPRADLTATSLIVSGGIIAAMLAHYHALMFLDAAHPILAVANPFGLTWLSGL	60		
QY	61	FIHADAMGSSVVGCRPRANAAGCOLVLMLYAGFSWRKMTVKSHMAHHNAGTDPPDDH	120		
DB	61	FIHADAMGSSVVGCRPRANAAGCOLVLMLYAGFSWRKMTVKSHMAHHNAGTDPPDDH	120		
QY	121	GGPRWTAARFGTYFGMRGELLRYUVTVTYVLMIGDRMMYVFWPFLBSILASIQLPVGGI	180		
DB	121	GGPRWTAARFGTYFGMRGELLRYUVTVTYVLMIGDRMMYVFWPFLBSILASIQLPVGGI	180		
QY	181	WLPHRPGDAEPDDHNAASSRISDPVSLTLCFHFSGYHNHNHNPPTVPMWMLPSTRIGD	240		
DB	181	WLPHRPGDAEPDDHNAASSRISDPVSLTLCFHFSGYHNHNHNPPTVPMWMLPSTRIGD	240		
QY	241	TA 242			
DB	241	TA 242			

```

RESULT 4
US-10-997-844-33
; Sequence 33, Application US/10997844
; Publication No. US20050124033A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Pamela L
; APPLICANT: Bosak, Melissa
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; APPLICANT: Cheng, Qiong L
; TITLE OF INVENTION: Optimized Bacterial Host Strains of Methylomonas sp. 16A
; FILE REFERENCE: CL-2230 US NA
; CURRENT APPLICATION NUMBER: US/10/997,844
; CURRENT FILING DATE: 2004-11-24
; PRIOR APPLICATION NUMBER: US 60/527,063
; PRIOR FILING DATE: 2003-12-03
; PRIOR APPLICATION NUMBER: US 60/527,877
; PRIOR FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33
; LENGTH: 242
; TYPE: prt
; ORGANISM: Agrobacterium aurantiacum
US-10-997-844-33

```

Query Match	99.3%	Score 1331	DB 18	Length 242
Best Local Similarity	98.8%	Pred. No. 4.6e-127		
Matches	239	Conservative	2	Mismatches 1, Indels 0, Gaps 0

Cy	1	MSAAHLPRADITATSLIVSGGIIAAMLALHVAHLMFLDAAHPILIAVANFICLTITLSTGL	60
Db	1	MSAAHLPRADITATSLIVSGGIIAAMLALHVAHLMFLDAAHPILIAVANFICLTITLSTGL	60

Db 191 PRRHTDQPPADAHARSSGYPVLSLTCTCFHF-GRHHEHLLSPWRPWRL 239

RESULT 8

US-10-848-307-39

/ Sequence 39, Application US/10848307

/ Publication No. US20050003474A1

/ GENERAL INFORMATION:

/ APPLICANT: desouza, Mervyn L.

/ APPLICANT: Schroeder, William A.

/ APPLICANT: Kollmann, Sherry R.

/ APPLICANT: May, Colleen A.

/ TITLE OF INVENTION: Carotenoid Biosynthesis

/ FILE REFERENCE: CGL00/0243US02

/ CURRENT APPLICATION NUMBER: US/10/848,307

/ CURRENT FILING DATE: 2004-05-18

/ PRIOR APPLICATION NUMBER: US 10/466,656

/ PRIOR FILING DATE: 2003-07-18

/ PRIOR APPLICATION NUMBER: US 60/288,984

/ PRIOR FILING DATE: 2001-05-04

/ PRIOR APPLICATION NUMBER: US 60/264,329

/ PRIOR FILING DATE: 2001-01-26

/ NUMBER OF SEQ ID NOS: 47

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 39

/ LENGTH: 244

/ TYPE: PRT

/ ORGANISM: Brevundimonas aurantiaca

US-10-848-307-39

Query Match 42.1%; Score 564; DB 17; Length 244;

Best Local Similarity 48.7%; Pred. No. 5.7e-49;

Matches 112; Conservative 28; Mismatches 76; Indels 14; Gaps 5;

QY 12 TATSLIVSGGIIAALAHVHALMF-----LDAAHPIIAVANFLGTLVSLVGLFIAND 66

Db 15 TWIGTLACMIIVAGVAVHIVYGVYFRRMGPTLTIVIAPAIVAVQ---TWLSVGLPIVAND 70

QY 67 AHGSGVVRGRANAMGQLVLMIVYGFSGRMKIVGMHHBAGTDDPDPDHGP--- 123

Db 71 AMYGSIAFGRPRINAAVGLITGLYAGFFRDLKTAHHHAAAPGADDPDPAAPAPAF 130

QY 124 VMVYARFIGTYFGWREGLLPVIIVTVYALMLGDR-WMVVYFWPPLPSIIASIQLFVFGIWL 182

Db 131 LPWFILNFFRTYFGWRMAVLTALVIALFGLGARPANLLTFWAAPALISALQLETFGTWL 190

QY 183 PRRFGDAPPRDHNAARSSRISDPVSLITCFHFGYTHHHHLPYVWML 232

Db 191 PRRHTDQPPADAHARSSGYPVLSLTCTCFHF-GRHHEHLLSPWRPWRL 239

RESULT 9

US-09-791-687A-4

/ Sequence 4, Application US/09791687A

/ Patent No. US20020053096A1

/ GENERAL INFORMATION:

/ APPLICANT: Hirschberg, Joseph

/ APPLICANT: Lotan, Tamir

/ TITLE OF INVENTION: POLYNUCLEOTIDE MOLECULE FROM HAEMATOCOCCUS PLUVIALIS ENCODING A

/ TITLE OF INVENTION: POLYPEPTIDE HAVING A BETA-C-4-OXYGENASE ACTIVITY FOR BIOTECHNOLO

/ TITLE OF INVENTION: OF (3S,3'S) ASTAXANTHIN AND ITS SPECIFIC EXPRESSION IN CHROMOLA

/ FILE REFERENCE: 01/21600

/ CURRENT APPLICATION NUMBER: US/09/791,687A

/ CURRENT FILING DATE: 2001-02-26

/ PRIOR APPLICATION NUMBER: US 09/259,294

/ PRIOR FILING DATE: 1999-03-01

/ NUMBER OF SEQ ID NOS: 4

/ SOFTWARE: PatentIn version 3.0

/ SEQ ID NO 4

/ LENGTH: 329

/ TYPE: PRT

/ ORGANISM: Haematococcus pluvialis

US-09-791-687A-4

Query Match 26.4%; Score 354.5; DB 9; Length 329;

Best Local Similarity 31.5%; Pred. No. 1.7e-27;

Matches 84; Conservative 44; Mismatches 104; Indels 35; Gaps 7;

QY 2 SAHALPKADLTATSLIVSGGIIAAMLAHVHALMFL-----DAAAHPIIA 46

Db 56 NAYKRPSPDGTGKITMALR--VIGSWAAVFLHAIFQIKLPTSLDQLHMLPVSDATQOLVSG 113

QY 47 VAN-----FLGLTWLSVGLFIANDMHGSSVYGRRAAAMQGLVLMIVYAGSWMK 99

Db 114 TSSLIDIVVFVFLVLEFVLTGFLPTTHDMHGITAMRNQLNDFGRVCISLYAMFDYML 173

QY 100 IVKMAHRRHAG-TDDDPDPHGP--VRMYARFIGTYFGWREGLLPVIIVTVYALMLGD 156

Db 174 HRKWEHHNHGVEVKDPPDFRGNPGIVPWFASFSMSYMSWQPARLAWMTVYVQQLGAP 233

QY 157 RMWYVFWPPLPSIIASIQLFVFGIWLPHRP-----GHDAFPDRHNAARSSRISDPVSL 209

Db 234 MANLLVFMAAPILISAFRLFYFGTYMPKRPBGASGSSPAVMWWMKSRTSQASDLVSFL 293

QY 210 TCFHFGYTHHHHLPYVWMLPSTR 236

Db 294 TCYHF-DLHWBHHRWFPAPMWELPNCR 319

RESULT 10

US-10-310-154-466

/ Sequence 466, Application US/10310154

/ Publication No. US20030233670A1

/ GENERAL INFORMATION:

/ APPLICANT: Edgerton, Michael D

/ APPLICANT: Chomet, Paul S.

/ APPLICANT: Adams, Thomas H

/ APPLICANT: Ruff, Thomas G.

/ APPLICANT: Agarwal, Ameeta K.

/ APPLICANT: Ahrens, Jeffrey E.

/ APPLICANT: Ball, James A.

/ APPLICANT: Banu, G.

/ APPLICANT: Bell, Erin

/ APPLICANT: Boddupalli, Raghava

/ APPLICANT: Delkman, Jill

/ APPLICANT: Deng, Molian

/ APPLICANT: Dong, Jinzhao

/ APPLICANT: Duff, Stephen M.

/ APPLICANT: Galligan, Meghan M.

/ APPLICANT: Hinchey, Brenda S.

/ APPLICANT: Huang, Shihshieh

/ APPLICANT: Johnson, G. Richard

/ APPLICANT: Jung, Vincent

/ APPLICANT: Kretzmer, Keith A

/ APPLICANT: Laccetti, Lucille B.

/ APPLICANT: Lai, Chao-Qiang

/ APPLICANT: Lee, Gary

/ APPLICANT: Lin, Jie-Yi

/ APPLICANT: Liu, Jingdong

/ APPLICANT: Lu, Bin

/ APPLICANT: Luechy, Michael M.

/ APPLICANT: Lund, Adrian

/ APPLICANT: Madson, Linda L.

/ APPLICANT: Malloy, Kathleen A.

/ APPLICANT: McKiel, Christine L.

/ APPLICANT: Miller, Philip W.

/ APPLICANT: Padmavathi, Manchikanti

/ APPLICANT: Parnell, Laurence D.

/ APPLICANT: Start, William G.

/ APPLICANT: Tennsen, Dan

/ APPLICANT: Vidya, K.R.

/ APPLICANT: Wang, Haiyun

/ APPLICANT: Xin, Zhaoguo

/ APPLICANT: Xu, Nanfei

/ APPLICANT: Yang, Chunzhi

```
APPLICANT: Zeng, Xiaoping
APPLICANT: Zhang, Qiang
APPLICANT: Zhao, Yajuan
APPLICANT: Zhou, Li
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
FILE REFERENCE: 38-15(52796)B
CURRENT APPLICATION NUMBER: US/10/310,154
CURRENT FILING DATE: 2002-12-04
PRIOR APPLICATION NUMBER: 60/337,358
PRIOR FILING DATE: 2001-12-04
NUMBER OF SEQ ID NOS: 736
SEQ ID NO 466
LENGTH: 453
TYPE: PRT
ORGANISM: Glycine max
US-10-310-154-466
```

```
Query Match      10.9%; Score 146; DB 15; Length 453;
Best Local Similarity 25.2%; Pred. No. 4.1e-06;
Matches 75; Conservative 34; Mismatches 95; Indels 94; Gaps 17;
```

```
QY 4 HALPKADLTATSLIVSGGIIAAMLAHVHAFDAAHPILAVANFL--GLTWLSVG-- 59
DB 117 HCMVKDPLKSMYSYVVDVIAVFGIA-----SAAAY-----LNNMLVWPLYMAAQGTM 163
QY 60 ---LPIAHADAMHGSVVPGRPRANAAMGOL---VIMLYAGFSRKMIVKMAHHRHAGT 112
DB 164 FMALEFVLGHDCGHS--FSNNPKLNSVVGHLHSSILVPYHG--WR--ISHRTTHQHGHG 217
QY 113 DDDPDPDHGCP-----VRWYARF-----IGTYFGMRGGLL 143
DB 218 VENDESMHPLPEKLRSLDTVTMRMLRFTAPFLAVPYVLMGRSPKTSHPDSSDLFV 277
QY 144 P-----VIVTV--YALMTGDRMYYVFWPLPSILASIOLE-----VFGIWL P----- 183
DB 278 PNERKDVITSTACWAMAMLG-----LVGLGFVWGPIQLLKLYGVPYVIFVWMLDLVTVL 331
QY 184 HRPQH-DAPPDRHNAASSRISDPVSL-----TCFPGYVHHHHHLPVPMWR 232
DB 332 HHGHEDKLPWYRGKEMSYLRGGLTTLDRDYGILNNIHHDIGTHVTHLFPQIPHYHL 389
```

RESULT 11

```
US-10-732-923-359
Sequence 359, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 359
LENGTH: 453
TYPE: PRT
ORGANISM: Glycine max
US-10-732-923-359
```

```
Query Match      10.9%; Score 146; DB 17; Length 453;
Best Local Similarity 25.2%; Pred. No. 4.1e-06;
Matches 75; Conservative 34; Mismatches 95; Indels 94; Gaps 17;
```

```
QY 4 HALPKADLTATSLIVSGGIIAAMLAHVHAFDAAHPILAVANFL--GLTWLSVG-- 59
DB 117 HCMVKDPLKSMYSYVVDVIAVFGIA-----SAAAY-----LNNMLVWPLYMAAQGTM 163
QY 60 ---LPIAHADAMHGSVVPGRPRANAAMGOL---VIMLYAGFSRKMIVKMAHHRHAGT 112
DB 164 FMALEFVLGHDCGHS--FSNNPKLNSVVGHLHSSILVPYHG--WR--ISHRTTHQHGHG 217
```

```
QY 113 DDDPDPDHGCP-----VRWYARF-----IGTYFGMRGGLL 143
DB 218 VENDESMHPLPEKLRSLDTVTMRMLRFTAPFLAVPYVLMGRSPKTSHPDSSDLFV 277
QY 144 P-----VIVTV--YALMTGDRMYYVFWPLPSILASIOLE-----VFGIWL P----- 183
DB 278 PNERKDVITSTACWAMAMLG-----LVGLGFVWGPIQLLKLYGVPYVIFVWMLDLVTVL 331
QY 184 HRPQH-DAPPDRHNAASSRISDPVSL-----TCFPGYVHHHHHLPVPMWR 232
DB 332 HHGHEDKLPWYRGKEMSYLRGGLTTLDRDYGILNNIHHDIGTHVTHLFPQIPHYHL 389
```

RESULT 12

```
US-10-732-923-4850
Sequence 4850, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 4850
LENGTH: 454
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(454)
OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-4850
```

```
Query Match      10.9%; Score 146; DB 17; Length 454;
Best Local Similarity 25.2%; Pred. No. 4.1e-06;
Matches 75; Conservative 34; Mismatches 95; Indels 94; Gaps 17;
```

```
QY 4 HALPKADLTATSLIVSGGIIAAMLAHVHAFDAAHPILAVANFL--GLTWLSVG-- 59
DB 118 HCMVKDPLKSMYSYVVDVIAVFGIA-----SAAAY-----LNNMLVWPLYMAAQGTM 164
QY 60 ---LPIAHADAMHGSVVPGRPRANAAMGOL---VIMLYAGFSRKMIVKMAHHRHAGT 112
DB 165 FMALEFVLGHDCGHS--FSNNPKLNSVVGHLHSSILVPYHG--WR--ISHRTTHQHGHG 218
QY 113 DDDPDPDHGCP-----VRWYARF-----IGTYFGMRGGLL 143
DB 219 VENDESMHPLPEKLRSLDTVTMRMLRFTAPFLAVPYVLMGRSPKTSHPDSSDLFV 278
QY 144 P-----VIVTV--YALMTGDRMYYVFWPLPSILASIOLE-----VFGIWL P----- 183
DB 279 PNERKDVITSTACWAMAMLG-----LVGLGFVWGPIQLLKLYGVPYVIFVWMLDLVTVL 332
QY 184 HRPQH-DAPPDRHNAASSRISDPVSL-----TCFPGYVHHHHHLPVPMWR 232
DB 333 HHGHEDKLPWYRGKEMSYLRGGLTTLDRDYGILNNIHHDIGTHVTHLFPQIPHYHL 390
```

RESULT 13

```
US-10-732-923-4730
Sequence 4730, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
```

NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 4730
LENGTH: 363
TYPE: PRT
ORGANISM: Rhodopseudomonas palustris
US-10-732-923-4730

Query Match 10.8%; Score 145; DB 17; Length 363;
Best Local Similarity 24.1%; Pred. No. 4e-06;
Matches 71; Conservative 39; Mismatches 84; Indels 100; Gaps 20;

QY 9 ADLITSLIVSGGIIAMTLAHVHALMFLDAAHP---ILAV--ANFGLTWSVGLFI 62
DB 44 AELITLAL---PLVALM-----SAAMFYSGHMAASLLIAIPAAGFL-----LRLFM 88
QY 63 IAHDMHGSVNPGRPRAN---AAMGQVLMLYAGSGWKMKVKKMAHHRHAGTDD--- 115
DB 89 ICHDCGHGAFFSKR-QANDWGRALGVITLTPD--CWR--QATHTATTGNLDRGV 142
QY 116 PPDFD---HGGPVRYARF-----IGTVF-----GMR--- 138
DB 143 GULDITVRESELSWGLKRLVKNPLVMFGLGPAVYFLQGRLPVGLMRDGRPMWS 202
QY 139 ---EGLIPVITYVYALMLGDRMKYVFWPLPSIIASIQLFVGIWL---PHRPGHDAF 191
DB 203 AMATNAALIGIVTL-TWFIGIKAFILVHLPIMLLAATA-----GWLIFYVOHQ--FETT 254
QY 192 PRRHARS-----RISDPVSLITGCFHGGYHNNHHPVPMWLP 233
DB 255 TWEHNRWSLHQALYGSYDLPAPLWFTANI--GMHVNHLMSRIPLYRLP 306

RESULT 14

US-10-369-493-15558
Sequence 15558, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiandeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15558
LENGTH: 268
TYPE: PRT
ORGANISM: Xanthomonas campestris
US-10-369-493-15558

Query Match 10.7%; Score 144; DB 15; Length 268;
Best Local Similarity 24.5%; Pred. No. 3.6e-06;
Matches 66; Conservative 44; Mismatches 79; Indels 80; Gaps 19;

QY 26 WLALHVALMFLDAAHPILAVANFLGLTWSVGLFIIAHDMHGSVNP---RPRANA 82
DB 1 WL---RALPFI--ALHLACFAVFWGVSWPAVGMAVALY-ALRMPALTGFYHRYFSHRA 53
QY 83 M---GQVLMLYAGFS-----WRKMIYKMAHHRHAGTDDPD-----FDHGP 123
DB 54 FKTSRVVOVFPAIATCVQGRPLW--WAHHRNHRHTDGTGRDPSPAQHGFWMSHTG- 110
QY 124 VRWY-----ARFITYFGMRBGLLPVITYVYALMLGDRM----- 159
DB 111 --WFLTPNFRDWEIVIDLRFTELRDLRFDIWPLVIALGLYLGD-WLAAAAAGLQ 167
QY 160 ---YVFWPLPSIIASIQLF--VFGI-WLPHRPGHDAFPDRHARSRIIDPVSILTCF 212

DB 168 TNGAQLVWGF--VLSTVALFHAFTIINSLAHRRGSRFRDTRDSRNWL---IALLT-- 220
QY 213 HFG--GYHHEHHLHP-----TVPWRLPST 235
DB 221 -FGBGMHNNHHPFGSARQGVWMEYDMT 248

RESULT 15

US-10-369-493-15931
Sequence 15931, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiandeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15931
LENGTH: 280
TYPE: PRT
ORGANISM: Xanthomonas campestris
US-10-369-493-15931

Query Match 10.7%; Score 144; DB 15; Length 280;
Best Local Similarity 24.5%; Pred. No. 3.8e-06;
Matches 66; Conservative 44; Mismatches 79; Indels 80; Gaps 19;

QY 26 WLALHVALMFLDAAHPILAVANFLGLTWSVGLFIIAHDMHGSVNP---RPRANA 82
DB 9 WL---RALPFI--ALHLACFAVFWGVSWPAVGMAVALY-ALRMPALTGFYHRYFSHRA 61
QY 83 M---GQVLMLYAGFS-----WRKMIYKMAHHRHAGTDDPD-----FDHGP 123
DB 62 FKTSRVVOVFPAIATCVQGRPLW--WAHHRNHRHTDGTGRDPSPAQHGFWMSHTG- 118
QY 124 VRWY-----ARFITYFGMRBGLLPVITYVYALMLGDRM----- 159
DB 119 --WFLTPNFRDWEIVIDLRFTELRDLRFDIWPLVIALGLYLGD-WLAAAAAGLQ 175
QY 160 ---YVFWPLPSIIASIQLF--VFGI-WLPHRPGHDAFPDRHARSRIIDPVSILTCF 212
DB 176 TNGAQLVWGF--VLSTVALFHAFTIINSLAHRRGSRFRDTRDSRNWL---IALLT-- 228
QY 213 HFG--GYHHEHHLHP-----TVPWRLPST 235
DB 229 -FGBGMHNNHHPFGSARQGVWMEYDMT 256

Search completed: June 17, 2005, 19:55:53
Job time : 83.6058 secs

GenCore version 5.1.6,
Copyright (c) 1993 - 2005 CompuGen Inc.

OM protein - protein search, using SW model

Run on: June 17, 2005, 19:18:35 ; Search time 20.9423 Seconds

(without alignments)
1111.838 Million cell updates/sec

Title: US-10-695-980-32

Perfect score: 1341

Sequence: 1 MSAAHLPRKADLTATSLIVSG.....LHPTVPMWRLPSTRKGDPA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	407.5	30.4	258	AF2204	beta-carotene keto
2	342.5	25.5	320	S65078	beta-carotene keto
3	190	14.2	685	C70678	probable electron
4	187.5	14.0	309	T31123	hypothetical prote
5	182	13.6	312	S77365	hypothetical prote
6	179.5	13.4	279	AB2307	hypothetical prote
7	158.5	11.8	345	T36617	probable Delta6 fa
8	152	11.3	369	S75038	hypothetical prote
9	143.5	10.7	329	D87673	conserved hypothet
10	135.5	10.1	350	S43772	phosphatidylcholin
11	135.5	10.1	350	AH2005	phosphatidylcholin
12	135.5	10.1	453	J02339	omega-3 fatty acid
13	135	10.1	424	J05891	omega-3 fatty acid
14	134	10.0	443	T01697	omega-3 fatty acid
15	131.5	9.8	398	T01696	omega-3 fatty acid
16	130.5	9.7	449	T50555	delta-8 sphingolip
17	130	9.7	449	AB4900	hypothetical prote
18	129.5	9.7	476	S52746	stearyl-CoA 9-des
19	128.5	9.6	383	T06238	omega-3 fatty acid
20	127	9.5	386	J02335	omega-3 fatty acid
21	126.5	9.4	387	T11959	fatty-acid desatur
22	126.5	9.4	351	S54259	Delta12 fatty acid
23	125.5	9.4	404	P00812	omega-3 fatty acid
24	125	9.3	377	J02337	omega-3 fatty acid
25	125	9.3	460	T10063	omega-3 fatty acid
26	123.5	9.2	381	T0323	probable omega-3 f
27	122.5	9.1	359	S52650	omega-3 fatty acid
28	122	9.1	355	D95270	probable fatty aci
29	120.5	9.0	446	J02336	omega-3 fatty acid

30	119	8.9	438	2	T15039	omega-3 fatty acid
31	118.5	8.8	380	2	T06235	omega-3 fatty acid
32	118.5	8.8	380	2	J02338	omega-3 fatty acid
33	118	8.8	383	1	A44227	omega-3 fatty acid
34	116	8.7	347	2	S43771	phosphatidylcholin
35	116	8.7	451	1	J06180	stearyl-CoA 9-des
36	115	8.6	285	2	AG2429	delta-8 desaturase
37	114	8.5	376	2	J07871	stearyl-CoA 9-des
38	113.5	8.5	351	2	S11519	phosphatidylcholin
39	112.5	8.4	349	2	S43770	phosphatidylcholin
40	112.5	8.4	441	2	T03029	omega-3 fatty acid
41	111	8.3	352	2	B63901	fatty-acid desatur
42	110.5	8.2	449	2	T47950	delta-8 sphingolip
43	109	8.1	360	2	T32554	hypothetical prote
44	108	8.1	332	2	A85062	probable fatty aci
45	108	8.1	379	2	J02555	omega-3 fatty acid

ALIGNMENTS

RESULT 1

AF2204 beta-carotene ketolase [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #ext_change 09-Jul-2004

C/Accession: AF2204

R/Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iritaghi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yaeuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AF2204

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-258 <KUR>

A/Cross-references: UNIPROT:O8YSA0; GB:BA000019; PIDN:BA074888.1; PID:G17132284; GSPDB:

C/Genetics:

A/Genes: alr3189

C/Superfamily: beta-carotene ketolase

C/Query Match

Best Local Similarity 30.4%; Score 407.5; DB 2; Length 258;

Matches 88; Conservative 39; Mismatches 87; Indels 13; Gaps 5;

QY 18 VSGIILAMWLAHVHLMPL-----DAAPHILAVANFLGLTWLSVGLFIADAMGS 71

DB 27 INKGIFACFLIFLWMSILSLIDSTSIHKSILGLIAMLWQTFYTGFIADAMHGV 86

QY 72 VVGRPRANAMGOLVMIYAGFSWRKMTVKMAHNRNAGTDDDDPFDGQYR--WTA 128

DB 87 VPKPKPRINNFGLKTLILYGLLIPYKDLKKMLHGHGTDLPDLYVNGHPQNFELWL 146

QY 129 RLITGVFGWR--GLLPTVTVVVALMLGDRMVVVFVWPLPSILASIOLEFVGIMLPHR 186

DB 147 HPMKSTWRTQFLGLVM-IFHGLKLVHLPENNLIIIFWIPILSSVQLFYGTFLPHK 205

QY 187 GHDAFPDRNARSRIIDPVSLITCFHFGVHNEHNLHPTVPMWRLP 233

DB 206 LGGGTNPICARSIPLEWMSFVTCYH-GYKHNHYQLPFWKLP 251

RESULT 2

S65078 beta-carotene ketolase - Haematococcus pluvialis

C/Species: Haematococcus pluvialis

C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #ext_change 09-Jul-2004

C/Accession: S65078

R/Kajiwara, S.; Kakizono, T.; Saito, T.; Kondo, K.; Ohtani, T.; Nishio, N.; Nagai, S.; M

Plant Mol. Biol. 29, 343-352, 1995

A/Title: Isolation and functional identification of a novel cDNA for astaxanthin biosynt

A:Reference number: S65078; MUID:96046752; PMID:7579184
A:Accession: S65078
A:Molecule type: mRNA
A:Residues: 1-320 <KAJ>
A:Cross-references: UNIPROT:Q39982; EMBL:D45881; NID:g1136638; PIDD:BA08300.1; PID:dl00
A:Experimental source: strain NIES-144
A:Note: It is uncertain whether Met-1, Met-8 or Met-33 is the initiator
C:Comment: This protein converts beta-carotene to canthaxanthin via echinonone. It may b
C:Superfamily: beta-carotene ketolase

Query Match 25.5%; Score 342.5; DB 2; Length 320;
Best Local Similarity 31.7%; Pred. No. 2.6e-23;
Matches 84; Conservative 39; Mismatches 107; Indels 35; Gaps 8;

QY 3 AAALPKADLTATSLVSGGIIAAMLAHVLA-----W-----FLDAA 40
DB 48 AKRPASDAKGTMTALT--ITGTWAVFLHAIFQRLPRTSMQMLPVPSEKTAOLGGS 105
QY 41 AHPILAANFLGLTLMVSGFLIAHDAMHGSVPRPANAAMGQLVLMYAGFSWRKMT 100
DB 106 SSLHIAAVFLVLEFLYGLFITTHDAMHGITALHRRQNDLGNICISLYAMFYSMLH 165
QY 101 VKNMAHHRAG-TDDDPDPDHGCP--VKRYAFICTYFGMRKGLLPVIVTYALMLGDR 157
DB 166 RKGWHEHHNHTGVEGKDPFKKNPGLVVPFASFMSSYMSLMQFARLAWVAVMQMLGAPM 225
QY 158 WMYVFWPPLPSILASIQLFVFGIMLPNRP-----GHDAFPDRHNASRSISDPVSLTLC 211
DB 226 ANLIVPMAAPILSAFLRYFYFTYTPHKRBPAGAGSQVMA-WPAKTSSEASDVNSFLTC 284
QY 212 FHFGYVHHHHLPTVPMWRPLSTR 236
DB 285 YHF-DLHWEHHRWPPAPWMLPHCR 308

RESULT 3

C70678
Probable electron transfer protein - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 16-Aug-2004

C:Accession: C70678
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70678
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-685 <COJ>
A:Cross-references: UNIPROT:P71846; GB:Z820298; GB:AL123456; NID:g3261664; PIDD:CA05067.
A:Experimental source: strain H37Rv
A:Gene: fdx
C:Genetics:
C:Superfamily: Beta-carotene ketolase/phenylacetate metabolism ferredoxin reductase, etc
C:Keywords: 2Fe-2S; metalloprotein
F:617-671/Domain: ferredoxin [2Fe-2S] homology <FER>
F:632,637,640,670/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 14.2%; Score 190; DB 2; Length 685;
Best Local Similarity 26.4%; Pred. No. 2.7e-09;
Matches 73; Conservative 38; Mismatches 98; Indels 68; Gaps 16;

QY 5 ALPK-----ADLTA---TSLIVSGGIIAAMLAHVLAHVLAFLDAAPILAVANFLGLTW 55
DB 43 ALPTGIFATLTAFGSTTAYISGW-IPFWYTI-----PVNAAVTFV----- 84
QY 56 LSVGLFIADHMGSSVPRPANAAMGQLVLMYAG--FSGRKVIYVQMAHHRAGTD 113
DB 85 ---MFTVVDHSHVAISSIR-WVNGLFGRLA-WLFVGVAVPAPFGYTHIQHHRSHSD 138

QY 114 D-DDD--FDHGG---PVRW-YARFITYGMRGGLLPVIVT-----VYA 151
DB 139 EODPDTFASHGSLWVLPRLMSMVEFYIKYLLPRGRSRPVEVAETLVMTLFLTGLIVA 198
QY 152 LMDGRMMYVFWPPLPSILASIQLFVFGIMLPNRPBGHA-PPDHNNASSRIS-----DP 205
DB 199 IVTGFWFLAVFLIPORTIGTLVLAWMDWLPNHLGLEDTORSNNRYRATRNVRGAEWLF 258
QY 206 VSLITCFHFGYVHHHHLPTVPMWRPLSTRKGDPA 242
DB 259 VLISQ-----NYHLVNHLPSPVFPFARYLTRRRNEA 290

RESULT 4

T31123

hypothetical protein 15 - Sphingomonas aromaticivorans plasmid pNL1
C:Species: Sphingomonas aromaticivorans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T31123
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.D.; Sisk, E.C.; Sensen, C.W.; G
submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati

A:Reference number: Z20992
A:Accession: T31123
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-309 <ROM>
A:Cross-references: UNIPROT:O85832; EMBL:AF079317; NID:g3378261; PID:g3378264; PIDD:AAD0
C:Genetics: plasmid pNL1
A:Note: orf015
C:Superfamily: beta-carotene ketolase

Query Match 14.0%; Score 187.5; DB 2; Length 309;
Best Local Similarity 26.9%; Pred. No. 2e-09;
Matches 77; Conservative 34; Mismatches 88; Indels 67; Gaps 17;

QY 5 ALPKADLTATSLVSGGIIAAM-LALH-VHALWFLDAAPILAVANFLGLTWSVGLFI 62
DB 24 ALPTALLFVWSM---AGIATTWYALNLSMLPLM-----VGAISNGL-VTYL---LFS 68
QY 63 IAHDMHGSV--VGRPANAAMGQLVLMYAGFSWRKMTYKMAHHRHAGTDDPD--FD 119
DB 69 VIHDSHKSLSVSGVINSIGALIGFLFPPAPVWVLRWV--HNKGHTHTGPKDPDFE 126
QY 120 HGGPRWVARFITYFG-----W-----RGG--L-----LPVIVTYALMLGDRW 158
DB 127 HESP-WWQVPRWTFDFGAYIMYFVKGQVNRKKGTELVFYSLLVFLFAAIFYFGCW 185
QY 159 MYVFWPPLPSILASIQLFVFGIMLPNRPBGHAFPDRHNASRSISDPVSLTLCFHF 215
DB 186 ELFMFMFVPSRLTFLIAIVFVILPHNP-----AIIAODEDF-MATWRFGMEW 234
QY 216 -----GYHNEHHLPTVPMWR 232
DB 235 LNLPLVQVNYHLIHLWPEIRPFYRM 260

RESULT 5

S77365
hypothetical protein s11468 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S77365
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77365
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-312 <KAN>
 A:Cross-references: UNIPROT:P73428; EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BA1746
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Genetics:
 A:Start codon: GTG
 C:Superfamily: beta-carotene ketolase

Query Match 13.6%; Score 182; DB 2; Length 312;
 Best Local Similarity 28.4%; Pred. No. 6.3e-09;

Matches 66; Conservative 41; Mismatches 71; Indels 54; Gaps 16;

QY 44 ILAVALPLGL-----TWLSVGLFIIA-----HDAMGSAVPGPRANAMQO---L 86
 DB 45 LMLAVALPGVWQWGLPHMLCFSCSVLAHLGTVIHDAH-NAAHNTIINAVLGHGSL 103
 QY 87 VLMVYGFSGWRKMTVGMMAHHRHAGTDDDPDPH---GGPV-----RWYARPTGYFG- 136
 DB 104 ML-----GFAFPVFTRVHLCQHNAV-NDEPNDPDHFVSTGRLFLIARFPYHEI-FFPKR 157
 QY 137 --WREGLLP-----VIVTVYALMLGDRMWYVF-----WPLPSILASTOLFPVGIWLP 184
 DB 158 RLNRKYLELWELSRLLVLTFT--VFLGIHGFIFGFWNTWFEVALLVGLALGFEDLP 215
 QY 185 RPSGDAFPDRHNAASRISDPVSLTCTFHFG-GYHHEHHLPTVPMWRPLST 235
 DB 216 RP-----FOERNRKNARVY-PSPLMLWLFQGVYHLHLHMLPSIPMYQYQNT 262

RESULT 6
 AB2307
 hypothetical protein alr4009 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AB2307
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MIMD:21595285; PMID:11759840
 A:Accession: AB2307
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-279 <KIR>
 A:Cross-references: UNIPROT:Q8YQ27; GB:BA000019; PIDN:BA875708.1; PID:G17133144; GSPDB:C
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr4009
 C:Superfamily: beta-carotene ketolase

Query Match 13.4%; Score 179.5; DB 2; Length 279;
 Best Local Similarity 28.6%; Pred. No. 9.4e-09;

Matches 68; Conservative 39; Mismatches 90; Indels 41; Gaps 17;

QY 16 LIVSGIILAMTLAHVHMLFLDAAHPIILAVANFLGLTWSGLFIADAMHGSVVR 75
 DB 2 LIFSASVVMILISNFGWLM---QMPWLCFSTNTTALHCAGT---VINDACHQSAHRN 54
 QY 76 RPRANAMQO---LVLMLVAGFSWRKMTVGMMAHHRHAG-TDDDP--FDHGSPVWY-- 127
 DB 55 R-IINMMLGHGSLIL---AFAFPVFTRVHLCQHNAVHNPKDDPDHYVSTGRL-MLLA 108
 QY 128 ARTI--GTYFG---WREGLLP-----VIVTVYALMLGDRMW-YV-FWPLPSILAST 173
 DB 109 VRFVLYHEVFFQKRLNRKYLELWELSRLLVITVIVISVQYHFLGYILNWFPIAFIVGI 168
 QY 174 QLVVPGIWLPHRSGDAFPDRHNAASRISDPVSLTCTFHFG-GYHHEHHLPTVPMW 230
 DB 169 ALGLFDPYLRHP---FVERDRWKARVY-PGKLMLILMGQYHLHLHMLPSIPMY 221

RESULT 7

T36617
 probable Delta6 fatty acid desaturase (EC 1.14.99.-) SCH35.42c [similarity] - Streptomyces
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T36617
 R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A:Reference number: Z21610
 A:Accession: T36617
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-345 <OLI>
 A:Cross-references: UNIPROT:Q9X8W4; EMBL:AL078610; PIDN:CA844385.1; GSPDB:GN00070; SCOE
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEB:SCH35.42c
 C:Superfamily: fatty acid desaturase/sphingolipid desaturase
 C:Keywords: oxidoreductase; unsaturated fatty acid biosynthesis

Query Match 11.8%; Score 158.5; DB 2; Length 345;
 Best Local Similarity 27.2%; Pred. No. 8.9e-07;

Matches 74; Conservative 32; Mismatches 83; Indels 83; Gaps 18;

QY 18 VSGGIILAMLA-LHYVALFLDAAHPIILAVANFLGLTWSGLFIADAMHGSVPCR 76
 DB 44 VTGLVYAGMAAFVLVQASWV-----TLAIAFLAVMYGOVAL--VAHDMARQVF-RR 93
 QY 77 PRANAMQOVLMLVAGFS-----WRKMTVGMMAHHRHAGTDD-DPPF-----D 119
 DB 94 RRASELSGRI-----AGASIGMSYGMWOD--KHTRHHPNPTEDLDPIGDPLVWSPD 145
 QY 120 HGGPVWYVRFITGYGMEGLLPVIVTV-----YALMLGD 156
 DB 146 QARAATGLPRLGR--WQAFPLPRLTLEGNTLVYASGRANRRLKRALDGLLH 202
 QY 157 RMYTV--VWPLPSILA---SIQLEVFGLWT--PHRPGH-----DAFPD--RHNR 198
 DB 203 CAVYLFALPMVLPFGMAIAFLAVHQCFLGVYLSAFAFPHKGMPIITADRPDLRRQVL 262
 QY 199 SSRISDPVSLTCTFHFGYTH--EHHLHPTV 228
 DB 263 TERNVNG-GLFTDLALGSLHQLIENHLFSPSP 293

RESULT 8
 S75038
 hypothetical protein sll1611 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S75038
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 A:Reference number: S74322; MIMD:97061201; PMID:8905231
 A:Accession: S75038
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-369 <KAN>
 A:Cross-references: UNIPROT:P73843; EMBL:D90910; GB:AB001339; NID:G1652956; PIDN:BA1790
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Genetics:
 A:Start codon: GTG

Query Match 11.3%; Score 152; DB 2; Length 369;
 Best Local Similarity 26.4%; Pred. No. 3.6e-06;

Matches 77; Conservative 27; Mismatches 84; Indels 104; Gaps 19;

QY 16 LIVSG---GIAMT--ALHVA-----LM-FLDAAHPIILAVANFLGLTWSVGLFI 63
 DB 28 LVILGLIGIYVAGLTIARHLHQWPGQWMLFLPMN-----LIMGN-----SVTVFLFG 76

QY 64 AHDHMGSVVGRPRAN---AAMGQVLVLTLAGSVWRKMVYKMAAHNRHNAETDDDP--- 117
 Db 77 SHDLMHGSLVKRRSKYSLISLIG-LSLMMWPRSQMRSL--HNQVHNNTNLSLRDPRTY 133
 QY 118 -----FDHGPVR--WYARFIGT-----YFGWRREGILLPVI 146
 Db 134 LHEQRKTKGWKIHNLFPARSGEVLNTLMLFGMGSTAMGVNIPRNLSVLEFTFGNADPVRPA 193
 QY 147 VTVALMLGDRMM-----YVFWPRLPSILASIQLFVFGIWLPRRG----- 187
 Db 194 FTVAQRDRQRLWLELAIGAVNLSTLFYQLQRLPIILG---YFLPFIQLGAMGMFYIY 249
 QY 188 --HNAFPRNHAARSRISDPV-----SLLSCFNFG--GYNHNHNHLPV 227
 Db 250 TNNHACP-----MTDINDPLVNSVSLMRKPLFQSLHNFSYHNEHNPV 295

RESULT 9

conserved hypothetical protein CC3422 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 16-Aug-2004
C:Accession: D87673
R:NIEMMAN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HEIDELBERG, J.
B.; LABB, M.T.; DEBOY, R.T.; DODSON, R.J.; DURKIN, A.S.; GWINN, M.D.; HAFE, D.H.; KOLONOS,
N.; J.; ERMOLOVA, M.; WHITE, O.; SALZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASER, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: D87673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <STO>
A:Cross-references: UNIPROT:O9A2Y4; GB:AE005673; NID:g13425136; PIDN:AAK25384.1, GSPDB:C
C:Genetics:
A:Gene: CC3422
A:Superfamily: Fatty acid (acyl-CoA) desaturase

	Query March	10.7%	Score 143.5;	DB 2;	Length 329;
	Best Local Similarity	23.9%;	Pred. No. 1.9e-05;		
	Matches	64;	Conservative	31;	Mismatches 96; Indels 77; Gaps 11;
Qy	21	GIIAMALAHVALWFLDDAAHPILAVANFLGLTWLSGLEFIIAHDAMHGSVGEGRPRAN	80		
Db	49	GVIILAAGALFY-----AFNPPLTYVLAWNLIGARQLGAILIIMHEAANGSLIH-KLVN	100		
Qy	81	AAMGOLVLMVA---GFSGRKMTIVKMAHHHAAGDDDPEDDHGGPARYKWYARFI-----	131		
Db	101	DWLGE--WLCAPRGASLASRYPIHLSHKKTAQAQAEPPDLVLSAFPFTTKRSLRKKIR	157		
Qy	132	----GYFGMRBGLL-----PVI-----VTVALMGDR	157		
Db	158	DLTGTFPEFKRFPRLPGMKGAKTTPGPKGAIFAGEIARORRPFLMNLIGILVLSALGLW	217		
Qy	158	WMYVVFWMLPSILASIQLFVFGIWL-----RRGDAPPD--RHNAEGSRISDP	205		
Db	218	WAMPRLMWLP-----MATWPFVLTRLRNIAEAHALVAKDEBDPPRIH-AKITTNAMI	266		
Qy	206	VSLITCFHGSGYHHEHHLAPTVPWMWRP	233		
Db	267	ERALLIAPYYVNFAEHHEMFMTTPCWNLP	294		

RESULT 10

phosphatidylcholine desaturase (EC 1.3.1.35) - *Anabaena variabilis* [misidentified]
N:Alternate names: Delta-12 desaturase; fatty acid desaturase (EC 1.14.99.-)
C:Species: *Anabaena variabilis*
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S43772
R:Sakamoto, T.; Wada, H.; Nishida, I.; Ohmori, M.; Murata, N.
Plant Mol. Biol. 24, 643-650, 1994

A:Title: Identification of conserved domains in the Delta-12 desaturases of cyanobacteria
 A:Reference numbers: S43770, MUID:94207189, PMID:8155883
 A:Accession: S43772
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-250 <Sak>
 A:Cross-references: UNIPROT:Q44503; GB:D14581; NID:G493223; PIDN:BA03435.1; PID:G493225
 C:Superfamily: fatty acid (acyl-CoA) desaturase
 C:Keywords: oxidoreductase

[illegible]

RESULT 11

phosphatidylcholine desaturase [imported] - Nostoc sp. (strain PCC 7120)
 A:Species: Nostoc sp. PCC 7120
 C:Spec: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 16-Aug-2004
 C:Accession: AH2005
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:117159840
 A:Accession: AH2005
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-350 <KUR>
 A:Cross-references: UNIPROT:Q44503; GB:BA000019; PIDN:BA87964.1; PID:G17135418; GSPDB:G6C7Genetics:
 A:Experimental source: strain PCC 7120
 A:Gene: desA
 C:Superfamily: Fatty acid (acyl-CoA) desaturase

	Query Match	10.1%	Score 135.5;	DB 2;	Length 350;
	Best Local Similarity	22.1%;	Pred. No. 0.0001;		
	Matches	62;	Conservative	33;	Mismatches 93; Gaps 16;
QY	SLIVSGGIIAM--TALTHTALMPIDAAAPHLLAVANPLGLTWLSVGEIITAHDAHMGSVV	73			
Dd	SVMVG--LGWSTLAI---APWFL---LPILMIFTLTALT-----GFVLGHDCGRHSFA	94			
QY	PGRRPANAAMGOLTLW--LYAGFSWRKVIYKHNAHHRACTDDPDPEDHGPVR-----	125			
Dd	K-RRMVDVLVGIIMMFLVIPHSNR--IKHNHHHGTHKTLEDNAMH--PIRPEVFAS	148			
QY	-----WTAFIG----TYFGRE-----	144			
Dd	WGTRTGSAFKLEMRQRILMTWGVSGHWMAVVHFEDMKCFVKDQADVKLSVAVVVLFAAVALFP	208			

C;Superfamily: fatty acid (acyl-CoA) desaturase

QY 186 PGH-DAFPDRHNA^{SSRI}SDPVL-----TCFHGGYHNEHH^{LPT}VPWRL 232
||| : : : ||| : : :
DB 323 HGHE^{DL}LPWYRGK^{EW}SLRG^LLTLLDRD^YGWINN^IHH^DIG^{TV}IHH^LFPQ^IRYH^L 378

RESULT 15

969101

omega-3 fatty acid desaturase (EC 1.14.99.-) FAD8 - maize (fragment)

C;Species: Zea mays (maize)

C;Date: 19-Feb-1999 #sequence 19-Feb-1999 #text_change 09-Jul-2004

C;Accession: T01696

R;Berberich, T.; Harada, M.; Sugawara, K.; Kodama, H.; Iba, K.; Kusano, T.

Plant Mol. Biol. 36, 297-306, 1998

A; Title: Two maize genes encoding omega-3 fatty acid desaturase and their differential

A;Reference number: Z14400; MUID:98145435; PMID:9484441

A;Accession: T01696

A; Status: preliminary; translated from GB/EMBL/DBJ

A: molecule type: mRNA

A;Residues: 1-398 <BER>

A; Cross-references: UNIPROT:O24626; EMBL:D63953; NID:g2446995; PIDN:BAA22440.1; PID:g244

A; Experimental source: strain honey bantam

C;Genetics:

A;Gene: FALD

C;Superfamily: fatty acid (acyl-CoA) desaturase

C;Keywords: oxidoreductase

Query Match	9.8%;	Score 131.5;	DB 2;	Length 398;
-------------	-------	--------------	-------	-------------

Best Local Similarity 24.1%; Pred. No. 0.00027;

Matches 67; Conservative 30; Mismatches 98; Indels 83; Gaps 15;

QY 16 LIVSGGIIAAMWLALHVALWFLDAAPPIAAVANFLGLTWLSVGLFIIAHDMHGSVVPG 75

Db 78 VVVVLGLAAARLDSWLVPPLYWAAQGT-----MFW---ALFVLGHDCHGS-FSN 125

QY 76 RPRANAMGOL-----VLWLYAGFSWRKMI VKMAHHRHAGTDDDPDFDHG GPVRWY----- 1277

Db 126 NPKLNSVGHILHSSILVRYHG--WR--ISHRTHQNHGHEKDESWHPLPERLYKSLD 180

QY 128 -----ARFI-----GTyFGwREGILLP-----VIVTVyALMLGDRW 158

Db 181 FMTRKLRFTMPFPLLAFFPLYLFARSPGKSGSHFNPSSDLFOPNEKDIITSTA-----SW 235

OV 159 MYVV-----FWPLPSI-LASIOLEVEGILP-----HRPGH-DAFPDRHNARSSRI 202

Db 236 LAMGVTLA¹GLTFELMPVAM²KL³YGVPYFV⁴FAW⁵LD⁶MTY⁷LHHGHED⁸KL⁹PMYRGOEWS¹⁰YL¹¹ 295

203 SDPVSL-----TCFHGGYHEHHLLHPTVPMWRL 2322

Db 296 RGGLTTI.DRDYGI.TNTHDICTHYVTHI.EPOTPHYH. 333

Search completed: June 17, 2005, 19:38:05

Job time : 21.9423 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 17, 2005, 18:55:34 ; Search time 91.1378 Seconds

(without alignments)
1359.734 Million cell updates/sec

Title: US-10-695-980-32

Perfect score: 1341
Sequence: 1 MSNAHLPKADLTATSLIVSG.....LHPTVPMRLPSTRKGDRA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1332	99.3	242	09RLH7	09RLH7 paracoccus
2	1331	99.3	242	1 CRTM_AGRAU	P54927 paracoccus
3	1044	77.9	242	1 CRTM_AUCSP	044261 alcaligenes
4	600	44.7	258	2 09KIX0	09KIX0 bradyrhizob
5	570	42.5	244	2 08GCT5	08GCT5 brevundimon
6	424.5	31.7	257	2 07NJV4	07NJV4 gloeobacter
7	407.5	30.4	258	2 08YSAO	08YSAO anabaena sp
8	406.5	30.3	229	2 0847D1	0847D1 nodularia s
9	359	26.8	244	2 07U6H0	07U6H0 haematococc
10	354.5	26.4	339	2 023973	023973 haematococc
11	350.5	26.1	320	1 06J3N5	06J3N5 haematococc
12	342.5	25.5	320	1 CRTM_HAEPL	039982 haematococc
13	300.5	22.4	276	2 08LJQ2	08LJQ2 haematococc
14	194.5	14.5	344	2 07UDV2	07UDV2 prochlorococ
15	190	14.2	672	2 07D5A7	07D5A7 mycobacteri
16	190	14.2	685	2 P71846	P71846 mycobacteri
17	190	14.2	685	2 07TW80	07TW80 mycobacteri
18	189	14.1	343	2 07U9G7	07U9G7 synecococc
19	187.5	14.0	309	2 08S832	08S832 sphingomonas
20	187	13.9	325	2 08DHP5	08DHP5 synecococc
21	182	13.6	312	2 P73428	P73428 synecococc
22	179.5	13.4	279	2 08YQ27	08YQ27 anabaena sp
23	167	12.5	346	2 07V4X0	07V4X0 prochlorococ
24	159.5	11.9	283	2 07NCW0	07NCW0 gloeobacter
25	158.5	11.8	345	2 09X8M4	09X8M4 streptomyc
26	152	11.3	369	2 P73843	P73843 synecocyst
27	148	11.0	379	2 06LJB6	06LJB6 photobacter
28	146.5	10.9	340	2 098120	098120 rhizobium 1
29	145	10.8	355	2 06N9E6	06N9E6 rhodospseudo
30	143.5	10.7	329	2 09A2Y4	09A2Y4 caulobacter
31	143.5	10.7	493	2 089T03	089T03 bradyrhizob

32	143	10.7	327	2	052580	052580 pseudomonas
33	141	10.5	584	2	052468	052468 pseudomonas
34	139	10.4	298	2	08PMP8	08PMP8 xanthomonas
35	139	10.4	301	2	08XRN9	08XRN9 ralsitonia s
36	139	10.4	341	2	08GGR8	08GGR8 streptomyc
37	139	10.4	350	2	0704F1	0704F1 noscocc sp.
38	138.5	10.3	465	2	07S1A7	07S1A7 neurospora
39	138	10.3	349	2	08PAZ4	08PAZ4 xanthomonas
40	136	10.1	320	2	06N921	06N921 rhodospseudo
41	135.5	10.1	350	2	079F72	079F72 anabaena va
42	135.5	10.1	350	2	044503	044503 anabaena sp
43	135.5	10.1	453	1	FD3C_SOYBN	FD3C_SOYBN max
44	135	10.1	424	2	048663	048663 chlamydomon
45	134	10.0	327	2	0828S4	0828S4 nitrosomona

ALIGNMENTS

RESULT 1

ID	09RLH7	PRELIMINARY;	PRT;	242 AA.
AC	09RLH7			
DT	01-MAY-2000 (TREMblrel. 13, Created)			
DT	01-MAY-2000 (TREMblrel. 13, Last sequence update)			
DT	01-MAR-2004 (TREMblrel. 26, Last annotation update)			
DE	Beta-carotene C-4-oxygenase (Ketolase).			
GN	Name=CTW;			
OS	Paracoccus marcusii.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;			
OC	Rhodobacteriaceae; Paracoccus.			
OX	NCBI_TaxId=59779;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=WH1;			
RA	Harker M., Hirschberg J.,			
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; Y15112; CAB56059.1; -			
DR	GO; GO:0016701; F:oxidoreductase activity, acting on single d. . .; IEA.			
DR	GO; GO:0016119; F:carotene metabolism; IEA.			
DR	InterPro; IPR011393; Carotene_ketolas.			
DR	InterPro; IPR010257; FA_desat.			
DR	InterPro; IPR010257; FA_desat sub.			
DR	Pfam; PF00487; FA_desaturase; 1.			
DR	PIRSF; PIRSF027840; Carotene_ketolas; 1.			
DR	ProDom; PD001061; FA_desat sub; 1.			
SQ	SEQUENCE 242 AA; 27114 NM; 5D98AD16412416C2 CRC64;			

Query Match 99.3%; Score 1332; DB 2; Length 242;

Best local Similarity 99.2%; Pred. No. 2,3e-110;

Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	MSNAHLPKADLTATSLIVSGGIIAAMLAHVAFWLDAAHPILAANFLGLTWLSVGL	60
DB	1	MSNAHLPKADLTATSLIVSGGIIAAMLAHVAFWLDAAHPILAANFLGLTWLSVGL	60
QY	61	FIIAHDAMGVSVPGRPRANAAMGQVLMVYGFSPKRVYGMHHRAGTDDDDPPH	120
DB	61	FIIAHDAMGVSVPGRPRANAAMGQVLMVYGFSPKRVYGMHHRAGTDDDDPPH	120
QY	121	GGPVRWYARFICGYFGMRGGLLPVIVTVYVPMKGDWYVFWPLPSILASIQLFVFT	180
DB	121	GGPVRWYARFICGYFGMRGGLLPVIVTVYVPMKGDWYVFWPLPSILASIQLFVFT	180
QY	181	WLFHRPGHDAFPDRNANASSRISDPVSLTCTHFGGYTHHHLLPVPWRLPSTRTKD	240
DB	181	WLFHRPGHDAFPDRNANASSRISDPVSLTCTHFGGYTHHHLLPVPWRLPSTRTKD	240
QY	241	TA 242	
DB	241	TA 242	

```

RESULT 2
CRTLW_AGRAU STANDARD; PRT; 242 AA.
ID CRTLW_AGRAU
AC P54972;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Beta-carotene ketolase (EC 1.13.-.-) (Beta-carotene oxygenase).
GN Name=crtlw;
OS Agrobacterium aurantiacum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=41455;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96062243; PubMed=7592436;
RA Misawa N., Satomi Y., Kondo K., Yokoyama A., Kajiwara S., Salto T.,
RA Ohtani T., Miki W.;
RT "Structure and functional analysis of a marine bacterial carotenoid
RT biosynthesis gene cluster and astaxanthin biosynthetic pathway
RT proposed at the gene level."
RT J. Bacteriol. 177:6575-6584(1995).
CC -!- FUNCTION: Converts beta-carotene to canthaxanthin via echinenone.
CC -!- PATHWAY: Carotenoid biosynthesis. Involved in astaxanthin
CC biosynthetic pathway.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D58420; BAA09591.1; -
CC InterPro: IPR005804; FA_desat.
CC InterPro: IPR010257; FA_desat_sub.
CC Pfam: PF00487; FA_desaturase; 1.
CC PIRSF: PIRSF027840; Carotene ketolase; 1.
CC Prodom: PD001081; FA_desat_sub; 1.
CC Carotenoid biosynthesis: Oxidoreductase.
CC SEQUENCE 242 AA; 27128 MW; 6AEB5BF60BE306B CRC64;

Query Match 99.3%; Score 1331; DB 1; Length 242;
Best Local Similarity 98.8%; Pred. No. 2.8e-110;
Matches 229; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSAAALPKADLTSTLSIVGGITAAWLAHVLAFLDAAPPIIAVANTLGLTWLSTGL 60
DB 1 MSAAALPKADLTSTLSIVGGITAAWLAHVLAFLDAAPPIIAVANTLGLTWLSTGL 60
QY 61 FTIAHDAMGSGVPPGRPRANAAGOLVLMVYAGFSRKKMIVKMAHRRHAGTDDDDPFDH 120
DB 61 FTIAHDAMGSGVPPGRPRANAAGOLVLMVYAGFSRKKMIVKMAHRRHAGTDDDDPFDH 120
QY 121 GSPVRYAFAFGTGYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGF 180
DB 121 GSPVRYAFAFGTGYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGF 180
QY 121 GSPVRYAFAFGTGYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGF 180
DB 121 GSPVRYAFAFGTGYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGF 180
QY 181 WLPHRPGHDAFPDRHNAARSRLSDPVSLITCFHFGYHHEHHLHPVPMWRLLPSTRTKGD 240
DB 181 WLPHRPGHDAFPDRHNAARSRLSDPVSLITCFHFGYHHEHHLHPVPMWRLLPSTRTKGD 240
QY 241 TA 242
DB 241 TA 242
DB 241 TA 242

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Beta-carotene ketolase (EC 1.13.-.-) (Beta-carotene oxygenase).
OS Alcaligenes sp.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Alcaligenes.
OX NCBI_TaxID=512;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95251715; PubMed=7733978;
RA Misawa N., Kajiwara S., Kondo K., Yokoyama A., Satomi Y., Salto T.,
RA Miki W., Ohtani T.;
RT "Canthaxanthin biosynthesis by the conversion of methylene to keto
RT groups in a hydrocarbon beta-carotene by a single gene."
RT Biochem. Biophys. Res. Commun. 209:867-876(1995).
CC -!- FUNCTION: Converts beta-carotene to canthaxanthin via echinenone.
CC -!- PATHWAY: Carotenoid biosynthesis. Involved in astaxanthin
CC biosynthetic pathway.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D58422; BAA09596.1; -
CC InterPro: IPR005804; FA_desat.
CC InterPro: IPR010257; FA_desat_sub.
CC Pfam: PF00487; FA_desaturase; 1.
CC PIRSF: PIRSF027840; Carotene ketolase; 1.
CC Prodom: PD001081; FA_desat_sub; 1.
CC Carotenoid biosynthesis: Oxidoreductase.
CC SEQUENCE 242 AA; 26939 MW; 077DA83FAAD14E4 CRC64;

Query Match 77.9%; Score 1044; DB 1; Length 242;
Best Local Similarity 77.5%; Pred. No. 9e-85;
Matches 179; Conservative 17; Mismatches 35; Indels 0; Gaps 0;

QY 12 TATSLIVGGITAAWLAHVLAFLDAAPPIIAVANTLGLTWLSTGLFTIAHDAMGSG 71
DB 12 TATSLIVGGITAAWLAHVLAFLDAAPPIIAVANTLGLTWLSTGLFTIAHDAMGSG 71
QY 72 VVPGPRANAAGOLVLMVYAGFSRKKMIVKMAHRRHAGTDDDDPFDHGGPVRYAFAFI 131
DB 72 VVPGPRANAAGOLVLMVYAGFSRKKMIVKMAHRRHAGTDDDDPFDHGGPVRYAFAFI 131
QY 132 GTYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGIWLPHRPGHDAF 191
DB 132 GTYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGIWLPHRPGHDAF 191
QY 132 STYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGIWLPHRPGHDAF 191
DB 132 STYFGMRGLLPVIVTYAALMDRMMYVFWPPLSTIASIQLFVFGIWLPHRPGHDAF 191
QY 192 PDRHNAARSRLSDPVSLITCFHFGYHHEHHLHPVPMWRLLPSTRTKGD 242
DB 192 PDRHNAARSRLSDPVSLITCFHFGYHHEHHLHPVPMWRLLPSTRTKGD 242

RESULT 4
Q9KIX0 PRELIMINARY; PRT; 258 AA.
ID Q9KIX0;
AC Q9KIX0;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Beta-carotene ketolase.
GN Name=crtlw;
OS Bradyrhizobium sp. ORS278.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=114615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORS278;

```

```

RX MEDLINE=20309720; PubMed=10851005;
RX DOI=10.1128/JB.182.13.3850-3853.2000;
RA Hannibal L., Lorguin J., Angles d'Aurelle N., Garcia N.,
RA Chantrel C., Maeson-Bolvin C., Dreyfus B., Giraud E.;
RT "Isolation and characterization of the cantaxanthin biosynthesis
RT genes from the photosynthetic bacterium Bradyrhizobium sp. strain
RT ORS278."
RL J. Bacteriol. 182:3850-3853 (2000).
DR EMBL: AF218415; MAF78203.1;
DR GO: GO:0016701; F:oxidoreductase activity, acting on single d. . .; IEA.
DR GO: GO:0016119; P:carotene metabolism; IEA.
DR InterPro: IPR011393; Carotene ketolase.
DR InterPro: IPR005804; Fa_desat.
DR InterPro: IPR010257; FA_desat_sub.
DR Pfam: PF00487; FA_desaturase.1.
DR PIRSF: PIRSF027840; Carotene ketolase; 1.
DR ProDom: PD001081; FA_desat_sub; 1.
SQ SEQUENCE 258 AA; 29736 MW; 4C0819F6C30465D0 CRC64;

Query Match 44.7%; Score 600; DB 2; Length 258;
Best Local Similarity 49.6%; Pred. No. 2.8e-45;
Matches 113; Conservative 35; Mismatches 74; Indels 6; Gaps 4;

QY 16 LTVSGGIIAAMLAHYALHFLDAAHPILAVANFLGL-TWLSVGLFTIADAMHGSVVP 74
DB 26 LTLAVITIAAMLVHVGAMFVPLTSLSLPALPLVLTQVLTWLVGLFTIADHMGSLVP 85
QY 75 GRPRANAMGQVLVLTLYAGFSWRKMTYKMAHHRHAGTDDDPDHCGR--VRMTARFT 131
DB 86 FKQVVRIRIGQLCLFLYAGFSFDALVVEHKKHRRHGTADDPDDEVPRPFHGMWASFP 145
QY 132 GTEFGREGILLVITYVYVYALMGDRMNY-VFMPLPSILASIQLFVGVLTLPHRGHDA 190
DB 146 LHTFGKQVAITIAVSLVQLVAVPLQNTLFLWALPGLISALQLTFPGYTLPHKPTOP 205
QY 191 FPDNRHARSRIISDPVSLTLCFFHGGYHHEHHLPTVPMWRLPSTRTK 238
DB 206 FADNRHARSFPAWLSLTLCFFH-GRHHEHHLHPDAPWRRLPBIKR 252

RESULT 5
Q8GCTS PRELIMINARY; PRT; 244 AA.
AC 08GCTS;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Beta-carotene C4 oxygenase.
GN Name=ctrlw;
OS Brevundimonas aurantiaca.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Brevundimonas.
OX NCBI_TaxID=74316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15266;
RA de Souza M.U., Kolimann S.R., Schroeder W.A.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY166610; AAN8630.2;
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro: IPR005804; Fa_desat.
DR InterPro: IPR010257; FA_desat_sub.
DR Pfam: PF00487; FA_desaturase; 1.
DR ProDom: PD001081; FA_desat_sub; 1.
SQ SEQUENCE 244 AA; 27261 MW; 17E57874433819DA CRC64;

Query Match 42.5%; Score 570; DB 2; Length 244;
Best Local Similarity 49.1%; Pred. No. 1.2e-42;
Matches 113; Conservative 27; Mismatches 76; Indels 14; Gaps 5;

QY 12 TATSLIVSGGIIAAMLAHYALHFLDAAHPILAVANFLGLTWLSVGLFTIADH 66
DB 15 TWIGLTLAGMTIVAGMVLHYGVYFHRMGVLTIVIAPAIYAVO---TWLSVGLFTIADH 70

```

```

QY 67 AMHGSVGRPRANAMGQVLVLTLYAGFSWRKMTYKMAHHRHAGTDDDPDHCGR--- 123
DB 71 AMHGSVGRPRANAMGQVLVLTLYAGFSWRKMTYKMAHHRHAGTDDDPDHCGR--- 130
QY 124 VRWYARFTITGFGREGILLPVITVYVYALMGDR-WMVVFWPLPSILASIQLFVGVLT 182
DB 131 LPMFLNFRFTYQWRMAVLTALVLTALFGLGARPNALLTFMAAPALISALQLFTGTWL 190
QY 183 PHRPGDAPDRHNAASSRIISDPVSLTLCFFHGGYHHEHHLPTVPMWR 232
DB 191 PRRHTDQPRADAHNAASSGVGPLSLTLCFFH-GRHHEHHLSPWRMRL 239

RESULT 6
Q7NOV4 PRELIMINARY; PRT; 257 AA.
AC Q7NOV4;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Beta-carotene ketolase.
GN Name=ctrlw; OrderedLocustNames=g111728;
OC Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneo T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Tanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids."
RL DNA Res. 10:137-145 (2003).
DR EMBL: AP006574; BAC89669.1;
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro: IPR005804; Fa_desat.
DR InterPro: IPR010257; FA_desat_sub.
DR Pfam: PF00487; FA_desaturase; 1.
DR ProDom: PD001081; FA_desat_sub; 1.
KW Complete proteome.
SQ SEQUENCE 257 AA; 29810 MW; 680EC9389FCC638C CRC64;

Query Match 31.7%; Score 424.5; DB 2; Length 257;
Best Local Similarity 38.0%; Pred. No. 1.1e-29;
Matches 95; Conservative 40; Mismatches 84; Indels 31; Gaps 9;

QY 14 TSLIVSGGII-----AAMLAHYALHFLDAAH---PI-LAVANFLGL---TWLS 57
DB 11 TSKRLAEVITHKNDSSGLWALVITIGLWIFFAALRLPIGELSLQAVIGVILTFPLH 70
QY 58 VGLFTIADAMHGSVGRPRANAMGQVLVLTLYAGFSWRKMTYKMAHHRHAGTDDDP 117
DB 71 TGLFTIADAMHRTVPAHNRINDWLGTAAGVGLYAMPYRELLIKQLHRRPATKDD 130
QY 118 F---DHGGRVYARFTITGFGREGILLPVITVYVYALMG-DRWV-----YVFWPL 166
DB 131 YHGBHSGFFQWYTKMKQY---ESRNTPELIIAGAAVVGCVTWLMGVPLVLAFLM 187
QY 167 PSTLASIQLFVGVITVPHRPGDAPDRHNAASSRIISDPVSLTLCFFHGGYHHEHHLPT 226
DB 188 PLVLSLQLFYFTYTLPHRPODGYNRRHRSNRLSPWSPVSCYH-GYHHEHHEYPL 246

QY 227 VPMWRLPSTR 236
DB 247 VPMWRLPSTR 256

RESULT 7
Q8YSAO

```

ID Q8YSA0 PRELIMINARY; PRT; 258 AA.
AC Q8YSA0;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Beta-carotene ketolase.
GN OrderedLocusNames=alr3189;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxId=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.,
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003592; BAB74888.1; -
DR PIR; AP2204; AP2204.
DR GO; GO:0016701; F:oxidoreductase activity, acting on single d. .; IEA.
DR GO; GO:0016119; P:carotene metabolism; IEA.
DR Pfam; PF00487; FA_desaturase; 1.
DR PIRSF; PIRSF027840; Carotene ketolase; 1.
DR Prodom; PD001081; FA_desat_sub; 1.
KW Complete proteome.
SQ SEQUENCE 258 AA; 30313 MW; 8F6C8E4ECF8E61B CRC64;

Query Match 30.4%; Score 407.5; DB 2; Length 258;
Best Local Similarity 38.8%; Pred. No. 3.6e-28;
Matches 88; Conservative 39; Mismatches 87; Indels 13; Gaps 5;

QY 18 VSGGIIAAMLAHVHLMFL-----DAAHPILAVANLGLTWISVGLFIADHAGS 71
DB 27 INKGIFACIFILFWAISLILSLIDTSIIHKSLGIALMOTFLYGLFIADHAGMV 86
QY 72 VVPGRPRAAAMGCVLWLVYAGFSWRKMKVKKMAHHRHAGTDDPDHFGVPR--WYA 128
DB 87 VVPGRPRIINFGKLLIILYGLPYKDLKKHMLHGHGTDIDPDYNGHQPFLMYL 146
QY 129 REIGTYFGWRE--GLLLPVITVYALMLGDRMYVFWPPLPSILASIQLFVFGIWLPHRP 186
DB 147 HEMKSYWMTQIFGLVM-IFHGLKULVHLPENNLIFFMWIPGISLVQFYGTFLPHK 205
QY 187 GHDAPDRNARSSRIDPVSLTCEHSGYHHEHHLHTVFWWRLP 233
DB 206 LEEGYTNPHCARISPLPLFWSPVTCYHF-GYHKEHHEYPQLPWWKLP 251

RESULT 8

ID Q847D1 PRELIMINARY; PRT; 229 AA.
AC Q847D1;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Putative beta-carotene ketolase.
OS Nodularia spumigena.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nodularia.
OX NCBI_TaxId=70799;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NSOR10;
RX MEDLINE=22549861; PubMed=12664164; DOI=10.1007/s00239-002-2415-0;
RA Moffett M.C., Neilan B.A.;
RT "Evolutionary affiliations within the superfamily of ketosynthases
reflect complex pathway associations."
RL J. Mol. Evol. 56:446-457(2003).
DR EMBL; AF210783; AA064399.1; -
DR GO; GO:0016701; F:oxidoreductase activity, acting on single d. .; IEA.

DR GO; GO:0016119; P:carotene metabolism; IEA.
DR InterPro; IPR011393; Carotene_ketolase.
DR InterPro; IPR005804; FA_desat.
DR InterPro; IPR010257; FA_desat_sub.
DR Pfam; PF00487; FA_desaturase; 1.
DR PIRSF; PIRSF027840; Carotene_ketolase; 1.
DR Prodom; PD001081; FA_desat_sub; 1.
SQ SEQUENCE 229 AA; 27832 MW; 71C9271BB371C95B CRC64;

Query Match 30.3%; Score 406.5; DB 2; Length 229;
Best Local Similarity 38.4%; Pred. No. 3.9e-28;
Matches 86; Conservative 41; Mismatches 76; Indels 21; Gaps 9;

QY 22 IIAAWLALVHLMFLDAAHPILAVANLGLTWISVGLFIADHAGSVPGRPRAN 80
DB 5 IISIV-AISLGLIYIDISQFKFWMLLPLIFWQGFYGLFITYADAHGVVFPNPKIN 63
QY 81 AAMGCVLWLVYAGFSWRKMKVKKMAHHRHAGTDDPDHFGVPR--WYAPRTGYFG 136
DB 64 HFIGSLCLFLYGLLPYQKLKKHMLHHNPASETPDF-HNGKQKNPFAMVLYFKRYS 122
QY 137 WREGLLPVITVYALMLGDRMY-----VFWPPLPSILASIQLFVFGIWLPHRPGHDAF 191
DB 123 MLIITLWITNLKTI-----WHPEDMNTYFWVVPISLSIQLFYGTFLPHSPVGCY 178
QY 192 PDRNARSSRIDPV--SLTTCFHSGYHHEHHLHTVFWWRLP 233
DB 179 KEPPH--RSQTISRPIWMSFITCYHF-GYHKEHHEYPVPMWQLP 219

RESULT 9

ID Q7U6H0 PRELIMINARY; PRT; 244 AA.
AC Q7U6H0;
DT 01-OCT-2003 (TREMblrel. 25, Created)
DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Possible beta-carotene ketolase.
GN OrderedLocusNames=SYMW1368;
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxId=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahmsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen B.E., McCarron J.,
RA Paulsen I.T., Dufrene A., Pattenisky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569692; CAB07883.1; -
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR005804; FA_desat.
DR InterPro; IPR010257; FA_desat_sub.
DR Pfam; PF00487; FA_desaturase; 1.
DR Prodom; PD001081; FA_desat_sub; 1.
KW Complete proteome.
SQ SEQUENCE 244 AA; 27434 MW; 77A71180F75875B7 CRC64;

Query Match 26.8%; Score 359; DB 2; Length 244;
Best Local Similarity 36.2%; Pred. No. 7e-24;
Matches 85; Conservative 41; Mismatches 97; Indels 12; Gaps 7;

QY 16 IIVSGGIIAAMLAHVHLMFLDAAHPILAVANLGLTWISVGLFIADHAGSVP 74
DB 13 IALAGLISAWL-LSLGSITPLDQTGGLIGSLILRALPLHGLFIADHAGSVALSV 71
QY 75 GRPRANAMGCVLWLVYAGFSWRKMKVKKMAHHRHAGTDDPDHFGVPR--HGCPVRYAFRI 131
DB 72 GHPGLNRWIGKYLIVVAGLSYERSRNRHRLHAPETFPDYPDRCNNNNILMDYVHF 131
QY 132 GTYFGWREGLLPVITVYALMLGD---RMYV-VFWPPLPSILASIQLFVFGIWLPHRP 186

```

Db      132 GNYLGNROLNLISCLWLAIIINGSDLPQIMHLLFSVLPILISSCOLFLVGTWLPFHR 191
Qy      187 GHDAFDRNRNARSISDPVSLTCTHFGGYHHEHLLHPVWRRLPSTRKCDT 241
      192 GATTRG-VTTRSLALPALSPACYNF-GYHREHHSBSPWFPOLPQINESFT 244

RESULT 10
ID      023973      PRELIMINARY;      PRT;      329 AA.
AC      023973;
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Beta-carotene C-4 oxygenase (Ketolase).
GN      Name=Cito;
OS      Haematococcus pluvialis.
OC      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC      Haematococcaceae; Haematococcus.
OX      NCBI_TaxID=44745;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=34/7;
RX      MEDLINE=97227932; PubMed=9119049; DOI=10.1016/S0014-5793(97)00110-5;
RA      Harter M., Hirschberg J.;
RT      "Biosynthesis of ketocarotenoids in transgenic cyanobacteria
      expressing the algal gene for 2-C-4-oxygenase, Cito.";
RL      FEBS Lett. 404:129-134(1997).
DR      EMBL; X86782; CAA60478.1;
DR      GO; GO:0016701; P:oxidoreductase activity, acting on single d. .; IEA.
DR      GO; GO:0016119; P:carotene metabolism; IEA.
DR      InterPro; IPR011393; Carotene_ketolase.
DR      InterPro; IPR005804; Pa_deeat.
DR      InterPro; IPR010257; PA_deeat_sub.
DR      Pfam; PF00487; PA_deeatutase; 1.
DR      PIRSF; PIRSF027840; Carotene_ketolase; 1.
DR      Prodom; PD001081; PA_deeat_sub; 1.
KW      Oxidoreductase.
SQ      SEQUENCE 329 AA; 37246 MW; 92027CC9AA4252FF CRC64;

Query Match      26.4%; Score 354.5; DB 2; Length 329;
Best Local Similarity 31.5%; Pred. No. 2.3e-23;
Matches 84; Conservative 44; Mismatches 104; Indels 35; Gaps 7;

Qy      2 SAHALPKADLTATSLVSGGIIAAMTALHYHAL-----DAAAPILA 46
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      56 NAYKPPSDTKGJITMALR--VIGSWAAVFLHAIFOIKLPTSLDQLHMLPVSDATQLVSG 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      47 VAN-----FLGLTWLVSGLFTIADAMHGSVVPGRPRANAMGOLVLMVYAGFSWRKY 99
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      114 TSSLDLIVVFFVLEFLYGLFTTTHDAMHGIIAMRRNDLGVCSILYAMFYNNML 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      100 IYKMAHHRHAG-TDDDPDFDHGSP--VRWYARFIGTFGWRBGLLPVIVTYALMLGD 156
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      174 HRKWEHNNHTGVGDPRFHKNPGIVPWFASFMSYSMWQFARLAWMTVMQLGAP 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      157 RMYTVVFWPLPSILASIQLFVFGIWLPHRP-----GHDAFDRNRNARSISDPVSL 209
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      234 MNTLVFMAAAPILISAFRLFYFGTYVPHKPEGASGSSPAVNMWKSRTSQASDLVSFL 293
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      210 TCFHFGYHHEHLLHPTVPMWRRLPSTR 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      294 TCTHF-DLHWEHHRWFPAPWMLPNCR 319
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
ID      06J3NS      PRELIMINARY;      PRT;      320 AA.
AC      06J3NS;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Beta-carotene ketolase.

```

```

GN      Name=bkt3;
OS      Haematococcus pluvialis.
OC      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC      Haematococcaceae; Haematococcus.
OX      NCBI_TaxID=44745;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NIES-144;
RA      Huang J.-C., Chen F., Sandmann G.;
RT      Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RL      EMBL; AY603347; AY3555.1;
DR      GO; GO:0016701; P:oxidoreductase activity, acting on single d. .; IEA.
DR      GO; GO:0016119; P:carotene metabolism; IEA.
DR      InterPro; IPR011393; Carotene_ketolase.
DR      InterPro; IPR005804; Pa_deeat.
DR      InterPro; IPR010257; PA_deeat_sub.
DR      Pfam; PF00487; PA_deeatutase; 1.
DR      PIRSF; PIRSF027840; Carotene_ketolase; 1.
DR      Prodom; PD001081; PA_deeat_sub; 1.
KW      Oxidoreductase.
SQ      SEQUENCE 320 AA; 35893 MW; 194EE30B2888FCDA CRC64;

Query Match      26.1%; Score 350.5; DB 2; Length 320;
Best Local Similarity 31.8%; Pred. No. 5.2e-23;
Matches 84; Conservative 39; Mismatches 108; Indels 33; Gaps 7;

Qy      3 AARLPKADLTATSLVSGGIIAAMTALHYHAL-----W-----FLDAA 40
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      48 AYKPPSDTKGJITMALR--IIGTWAVFLHAIFOIKLPTSMQDLMVPSAATQQLGGS 105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      41 AHPILVANFLGLTWLVSGLFTIADAMHGSVVPGRPRANAMGOLVLMVYAGFSWRKMI 100
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      106 SSIHRTAAVFVLEFLYGLFTTTHDAMHGIIAMRRNDLGNICISLYAMFYNNML 165
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      101 YKMAHHRHAG-TDDDPDFDHGSP--VRWYARFIGTFGWRBGLLPVIVTYALMLGDR 157
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      166 RKHWEHNNHTGVGDPRFHKNPGIVPWFASFMSYSMWQFARLAWMTVMQLGAP 225
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      158 RMYTVVFWPLPSILASIQLFVFGIWLPHRPHGHAFFPRH-----NARSISDPVSLTCT 212
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      226 ANILVFMMAAPILISAFRLFYFGTYVPHKPEGASGQVMAWFRKTSASDVMSPLTCY 285
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      213 HFGGYHHEHLLHPTVPMWRRLPSTR 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      286 HF-DLHWEHHRWFPAPWMLPNCR 308
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
ID      CRTW_HAEPL      STANDARD;      PRT;      320 AA.
AC      039982;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Beta-carotene ketolase (EC 1.13.-.-) (Beta-carotene oxygenase).
OS      Haematococcus pluvialis.
OC      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC      Haematococcaceae; Haematococcus.
OX      NCBI_TaxID=44745;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NIES-144;
RX      MEDLINE=96046752; PubMed=7579184;
RA      Kajiwara S., Kakizono T., Saito T., Kondo K., Ohtani T., Nishio N.,
      Nagai S., Misawa N.;
RT      Isolation and functional identification of a novel cDNA for
      astaxanthin biosynthesis from Haematococcus pluvialis, and astaxanthin
      synthesis in Escherichia coli ";
RL      Plant Mol. Biol. 29:343-352(1995).
CC      -1- FUNCTION: Converts beta-carotene to canthaxanthin via echinenone.
CC      -1- PATHWAY: Carotenoid biosynthesis. Involved in astaxanthin
      biosynthetic pathway.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: D45881; BAA08300.1; -.
CC PIR: S65078; S65078.
CC

DR InterPro: IPR005804; FA_desat.
DR InterPro: IPR010257; FA_desat_sub.
DR Pfam: PF00487; FA_desaturase; 1.
DR PIRSF: PIRSF027840; Carotene ketolase; 1.
DR ProDom: PD001081; FA_desat_sub; 1.
DR Carotenoid biosynthesis; Oxidoreductase.

KW Carotenoid biosynthesis; Oxidoreductase.
SQ SEQUENCE 320 AA; 35988 MW; 4FAD8669BD326527 CRC64;

Query Match 25.5%; Score 342.5; DB 1; Length 320;
Best Local Similarity 31.7%; Pred. No. 2.7e-22;
Matches 84; Conservative 39; Mismatches 107; Indels 35; Gaps 8;

QY 3 AHAAPKADLTATSLVSGGIIAAMLAHVAH-----W-----FLDAA 40
DB 48 AYKPPASDAKGIWMAIT--IGTWAVFLHAIQIRLPTSDQLHMLPVSEATQILGGS 105
QY 41 AHPILAVANFLGLTWLSVGLFTIAHDAMGSGVPGRPANAMGQVLVLTAGSWMKMT 100
DB 106 SSSLHIAAVFLVEFLYGLFTTTDHAMGTALRNOLDLGNICISLYAMPDYS 165
QY 101 VKHMAHHRHAG-TDDDPDFDHGSP--VRWYARFIGTYGCRGLLPVIVTVYALMLGDR 157
DB 166 RGHMEHNNHTGVBGKDPDFHKGNPGLVPWFASFMSYMSLMQFARLAWAVVMQMLGAPM 225
QY 158 WMYVVFVFWPLPSILASIQLFVFGIWLPHRP-----GHDAFPDRNARSRRISDPVSL 211
DB 226 ANLVFMAAAPTLSARFLFYFGIYLPKHPKBPAGASQVMA-WPRAKTSBADVMSPLTC 284
QY 212 FHHGGYHHEHHLHPVFWMLPSTR 236
DB 265 YHF-DLHMEHHRMPAPWQLPHCR 308

RESULT 13

Q8LJQ2 PRELIMINARY; PRT; 276 AA.

AC O8LJQ2: 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE BKT.
OS Haematococcus pluvialis.
OC Haematococcales; Chlorophyta; Chlorophyceae; Volvocales;
OC Haematococcaceae; Haematococcus.
OX NCBI_TaxID=44745;
RN [1]
RP SEQUENCE FROM N.A.
RA Teng C.Y., Qin S., Tseng C.K.;
RU Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF534876; FAN03484.1; -.
DR GO: GO:0016701; F:oxidoreductase activity, acting on single d. . .; IEA.
DR GO: GO:0016119; P:carotene metabolism; IEA.
DR InterPro: IPR011393; Carotene ketolase.
DR InterPro: IPR005804; FA_desat.
DR InterPro: IPR010257; FA_desat_sub.
DR Pfam: PF00487; FA_desaturase; 1.
DR PIRSF: PIRSF027840; Carotene ketolase; 1.
DR ProDom: PD001081; FA_desat_sub; 1.
KW Oxidoreductase.
SQ SEQUENCE 276 AA; 30854 MW; 8967DDCC07A0BB90 CRC64;

Query Match 22.4%; Score 300.5; DB 2; Length 276;
Best Local Similarity 30.2%; Pred. No. 1.2e-18;

Matches 81; Conservative 39; Mismatches 95; Indels 53; Gaps 11;

QY 3 AHAAPKADLTATSLVSGGIIAAMLAHVAH-----W-----FLDAA 40
DB 16 AYKPPASDAKGIWMAIT--IGTWAVFLHAIQIRLPTSDQLHMLPVSEATQILGGS 73
QY 41 AHPILAVANFLGLTWLSVGLFTIAHDAMGSGVPGRPANAMGQVLVLTAGSWMKMT 100
DB 74 SSSLHIAAVFLVEFLYGLFTTTDHAMGTALRNOLDLGNICISLYAMPDYS--- 130
QY 101 VKHMAHHRHAG-TDDDPDFDHGSP--VRWYARFIGTYGCRGLLPVIVTVYALML 153
DB 131 -MHEHNNHTGVBGKDPDFHKGNPGLVPWFASFMSYMSLMQFARLAWAVVMQMLGAPM 189
QY 154 LGDRMYVVFVFWPLPSILASIQLFVFGIWLPHRP-----GHDAFPDRNARSRRISDPVSL 208
DB 190 AN-----LLVFMMAAPTLSARFLFYFGIYLPKHPKBPAGASQVMSWPAKTSBADVMSF 245
QY 209 LTCFHGGYHHEHHLHPVFWMLPSTR 236
DB 246 LTCYHFDLP-----APWMLPHCR 264

RESULT 14

Q7VDV2 PRELIMINARY; PRT; 344 AA.

AC Q7VDV2: 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Beta-carotene hydroxylase (EC 1.14.13.-).
GN Name=crtr; OrderedLocustNames=Pro0266;
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAG / CCMF 1375 / SS120;
RA MEDLINE=22810154; PubMed=12917486; DOI=10.1073/pnas.1733211100;
RA Dufrene A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Ozta S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weisenbach J., Wincker P.,
RA Wolf Y.I., Hees W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal photosynthetic genome."
RT Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
DR EMBL: AB017161; AAP9312.1; -.
DR GO: GO:0016701; F:oxidoreductase activity, acting on single d. . .; IEA.
DR GO: GO:0016119; P:carotene metabolism; IEA.
DR InterPro: IPR011393; Carotene ketolase.
DR InterPro: IPR005804; FA_desat.
DR InterPro: IPR010257; FA_desat_sub.
DR Pfam: PF00487; FA_desaturase; 1.
DR PIRSF: PIRSF027840; Carotene ketolase; 1.
DR ProDom: PD001081; FA_desat_sub; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 344 AA; 40462 MW; 5D98CAEAF018813A CRC64;

Query Match 14.5%; Score 194.5; DB 2; Length 344;
Best Local Similarity 29.2%; Pred. No. 4.1e-09;
Matches 68; Conservative 38; Mismatches 88; Indels 39; Gaps 15;

QY 20 GGIITAMLAHVAHMLFIDAAHPIAVANFLGLTWLSVGLFTIAHDAMGSGVPGRP 78
DB 46 GGYIAIFLISW---QWYGVMPDLPLV-----VGLAFSLHMEGTIVHDACHKAHPNK-w 96
QY 79 ANAAGQLVLMVLYAGFSWMKMTVKMAHHRHAGTDDDPDFH-----GSPVWYAR---FI 131
DB 97 INQMGHGA--ILGFSFVFTRVHQLQHSNV-NDPKNDPRHIVSTFGPV-WLAPRRFY 153
QY 132 GTYFG-----WRE-----GLLPVIVTVYALMLGDRMYVVF--WPLPSILASIQLFVF 178


```

Db      154 HEFFPGRKLMRYELMGMGLERSIFITVLAGVHFNFMNVINYLMFGALMVGVTLAGIF 213
Qy      179 GYMLPRPGHDAPPDHNAARSSRISDPVSLITCFHGG-GYHHEHHLHPVPPW 230
Db      214 FVLPFRP-----FMANNKWNKSRVY-PSRYVMNLLMGQYHVLVHLMPSIPWP 261

```

RESULT 15

```

Q7D5A7 ID Q7D5A7 PRELIMINARY; PRT; 672 AA.
AC Q7D5A7;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, last annotation update)
DE Fatty acid desaturase, putative/ferredoxin reductase, electron
DE transfer component.
GN OrderedlocusNames=MT3658;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J.C., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Frazer C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
DR J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AE000516; AAK48018.1; -.
DR TIGR; MT3658; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006058; 2FE2S fd BS.
DR InterPro; IPR008333; FAD binding_6.
DR InterPro; IPR005804; Fd_desat.
DR InterPro; IPR010257; Fd_desat_sub.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR001709; FPN_cyt_redctse.
DR InterPro; IPR001433; Oxred_FAD/NAD(P).
DR Pfam; PF00970; FAD binding_6; 1.
DR Pfam; PF00487; Fd_desaturase; 1.
DR Pfam; PF00111; Fer2; 1.
DR Pfam; PF00175; NAD binding_1; 1.
DR PRINTS; PR00371; FPNCR.
DR PRODOM; PD001081; Fd_desat_sub; 1.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; UNKNOWN_1.
KW Iron; Iron-sulfur; Metal-binding.
SQ SEQUENCE 672 AA; 75225 MW; 6E5B47B6CD1AB57B CRC64;

```

Query Match 14.2%; Score 190; DB 2; Length 672;

Best Local Similarity 26.4%; Pred. No. 2e-08;

Matches 73; Conservative 38; Mismatches 98; Indels 68; Gaps 16;

```

Qy      5 ALPK-----ADLTG-----TSLIVSGGIIAAMLAHVHAWFLDAAHPITIAVNPGLTW 55
Db      30 ALPTIGIFLATLTAFAVSGTAYISGW-IPFWTYI-----PVNAAVTFV----- 71
Qy      56 LSYGLFLIADHAGSVVPRPRANAMGQLVLMVAG--FSMRKMTVKMAHHRAGTD 113
Db      72 ---MFTVVDASHVAISSIR-WVNGIFGRLA-WLTVGPVAVFAPAFGYIHIOHRRSND 125
Qy      114 D-DPD--FDHGG--PVRW-YARFIGTVFGREGILLPVIVT-----VVA 151
Db      126 EQDDPDPFASHGSLMWLPFRKSWVEFYIKYLLPRGRSPVIEVAETLVMMTLLFLTGLIVA 185

```

```

Qy      152 LMLGDRMMYVFWPPLPSILASIQLFVFGIWLPHRPGHDA-FPDRHNAARSSRIS-----DP 205
Db      186 IYTGAFMTLAIYFLLPQRIGLTVLAWFPMWLFHHGLEDTQRGRNRYRATNNRGAEWLF 245
Qy      206 VSLITCFHGGYHHEHHLHPVPMWRLPSTRTKGDTA 242
Db      246 VLLSQ-----NYHLVHLLPVSVPFYRYLRTWRNEBA 277

```

Search completed: June 17, 2005, 19:37:06
job time : 94.1378 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 15:45:26 ; Search time 477.072 Seconds
(without alignments)
9045.789 Million cell updates/sec

Title: US-10-695-980-31

Perfect score: 729

Sequence: 1 atgagcgcacatcgcctcgc.....ccaagg99gacacccgcatga 729

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	729	100.0	729	8	ACA99471
2	729	100.0	1261	2	AAV40170
3	727.4	99.8	729	2	AAV40147
4	698.6	95.8	729	12	ADO61160
5	698.6	95.8	729	12	ADO61077
6	698.6	95.8	729	12	ADO38246
7	698.6	95.8	729	13	ADR03861
8	698.6	95.8	729	13	ADR03941
9	697	95.6	729	12	ADO61156
10	697	95.6	729	12	ADO61073
11	697	95.6	729	12	ADO38242
12	697	95.6	729	13	ADR03857
13	697	95.6	729	13	ADR03937
14	611.8	83.9	639	2	AAV64074
15	611.8	83.9	639	2	AAV64074
16	446	61.2	1944	11	ADG96833
17	446	59.6	1966	11	ADG96933
18	434.2	59.6	747	2	AAV73184
19	434.2	59.6	902	2	AAV25068
20	434.2	59.6	902	6	ABA97360

21	432.4	59.3	729	2	AAQ99487
22	432.4	59.3	729	2	AAV84077
23	432.4	59.3	1631	12	ADO61158
24	432.4	59.3	1631	12	ADO61075
25	432.4	59.3	1631	12	ADO38244
26	432.4	59.3	1631	13	ADR03859
27	432.4	59.3	1631	13	ADR03939
28	383.8	52.6	726	2	AAV40179
29	217	29.8	776	12	ADO61164
30	217	29.8	776	12	ADO61081
31	217	29.8	776	12	ADO38250
32	217	29.8	776	13	ADR03865
33	217	29.8	776	13	ADR03945
34	217	29.8	777	10	ADC99012
35	195.4	26.8	735	8	ABT14221
36	193	26.5	332	2	AAV73205
37	136.6	18.7	250	2	AAV73204
38	111.4	15.3	208	2	AAV73206
39	82.6	11.3	942	2	AAV18013
40	82.6	11.3	1662	12	ADO61154
41	82.6	11.3	1662	12	ADO61071
42	82.6	11.3	1662	12	ADO38240
43	82.6	11.3	1662	13	ADR03855
44	82.6	11.3	1662	13	ADR03935
45	82.6	11.3	2543	4	AAV57334

ALIGNMENTS

RESULT 1
ID ACA99471 standard; DNA; 729 BP.
ACA99471;
28-JUL-2003 (first entry)
DNA encoding enzyme polypeptide #10.
Gene; d: hydroxymethylglutaryl coenzyme-A reductase; HMGCoA reductase;
isopentenyl diphosphate isomerase; HMGCoA synthase; mevalonate kinase;
phosphomevalonate kinase; diposphomevalonate decarboxylase; phycoene;
carotenoid-producing cell; isoprenoid compound; FPP synthase; lycopen;
farnesyl diphosphate synthase; carotenoid; beta-carotene; zeaxanthin;
kew cantilaxanthin; astaxanthin; adonixanthin; cryptoxanthin; echinenone;
adonitribin; C-4-isoprenoid compound; nutritional supplement;
food colourant; animal feed pigment.
OS Paracoccus carotinifaciens.
P-PSDB; ABU97244.
12-DEC-2002.
05-JUN-2002; 2002MO-EP006171.
06-JUN-2001; 2001US-0296299P.
(HOFF) ROGHE VITAMINS AG.
Berry A, Bretzel W, Huemelin M, Lopez-Ulibarri R, Mayer AF;
Yellisev A;
WPI; 2003-156856/15.
P-PSDB; ABU97244.
New isolated polypeptide used for isoprenoid production, includes amino
acid sequence having e.g. mevalonate kinase, phosphomevalonate kinase or
diphosphomevalonate decarboxylase activity.
Example 12; Page 294-295; 307pp; English.

CC The invention relates to an isolated polypeptide having the activity of
CC an enzyme such as hydroxymethylglutaryl Coenzyme-A (HMGCoA) reductase,
CC isopentenyl diphosphate isomerase, HMGCoA synthase, mevalonate kinase,
CC phosphomevalonate kinase or diphosphomevalonate decarboxylase. The
CC invention also relates to the production of a carotenoid-producing cell
CC and engineering a bacterium to produce an isoprenoid compound, comprising
CC culturing a parent bacterium in a medium permitting expression of an
CC isoprenoid compound and selecting a mutant bacterium from the culture
CC medium that produces 1.1-1000 times more isoprenoid compound than the
CC parent bacterium, introducing an expression vector comprising a *farnesyl*
CC diphosphate (FPP) synthase DNA fragment operably linked to an expression
CC control sequence into the mutant bacterium, and selecting a bacterium
CC that contains the expression vector and produces at least 1.1 times more
CC isoprenoid compound than the selected mutant. The cell is useful for
CC producing a carotenoid (phycoene, lycopene, beta-carotene, zeaxanthin,
CC canthaxanthin, astaxanthin, adonixanthin, cryptoxanthin, echinenone or
CC adonirubin) by culturing the cell under conditions permitting expression
CC of the polypeptide encoded by the polynucleotide sequence and isolating
CC the carotenoid from the cell or the medium of the cell. The carotenoids
CC (C-4- isoprenoid compounds) are used as nutritional supplements,
CC pharmaceuticals, food colourants and pigments for animal feeds. This
CC sequence represents DNA encoding a polypeptide of the invention
XX
XX
SQ Sequence 729 BP; 100 A; 244 C; 232 G; 153 T; 0 U; 0 Other;

Query Match 100.0%; Score 729; DB 8; Length 729;
Best Local Similarity 100.0%; Pred. No. 1.6e-153;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGCATGCGCTTCCGCAAGCGAGATGACCGGCAACAGTTTGGCTCTCGGAGC 60
DB 1 ATGAGCGCATGCGCTTCCGCAAGCGAGATGACCGGCAACAGTTTGGCTCTCGGAGC 60
QY 61 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 120
DB 61 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 120
QY 121 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 180
DB 121 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 180
QY 181 TTCATCATGCGCGATGACGCGATGATGCGGTGCTGTCGCGCGCGCGCGCGCAT 240
DB 181 TTCATCATGCGCGATGACGCGATGATGCGGTGCTGTCGCGCGCGCGCGCGCAT 240
QY 241 GGGGCGATGCGCGCGATGCTGCTGCGTGCATGCGCGATGCTGCGCGCGCGCGAGTGC 300
DB 241 GGGGCGATGCGCGCGATGCTGCTGCGTGCATGCGCGATGCTGCGCGCGCGCGAGTGC 300
QY 301 GTCAAGCATGCGCGCGATGCGCGATGCGCGAGACCGAGACCGACATTTTCGACAT 360
DB 301 GTCAAGCATGCGCGCGATGCGCGATGCGCGAGACCGAGACCGACATTTTCGACAT 360
QY 361 GGGGCGCGCGCGTGGCTGTAACCGCGCGCTTCATGCGCGCATTTTGGCGTGGCGAGAGG 420
DB 361 GGGGCGCGCGCGTGGCTGTAACCGCGCGCTTCATGCGCGCATTTTGGCGTGGCGAGAGG 420
QY 421 CTGCGTGCCTCCGCTATGCGTGCATGCGGTGCGTGCATGCGGTGCGTGCATGCGGTGCG 480
DB 421 CTGCGTGCCTCCGCTATGCGTGCATGCGGTGCGTGCATGCGGTGCGTGCATGCGGTGCG 480
QY 481 GTGATCTTCTGGCGCTTGGCGCTGCGATCCTGCGCTGCGATCAGCTGTTCTGTTCCGAGTC 540
DB 481 GTGATCTTCTGGCGCTTGGCGCTGCGATCCTGCGCTGCGATCAGCTGTTCTGTTCCGAGTC 540
QY 541 TGGCTGCGCGACCGCGCGCGCGCGACGAGCGGTTCCGCGACCGACCAATGCGCGGTGCG 600
DB 541 TGGCTGCGCGACCGCGCGCGCGCGACGAGCGGTTCCGCGACCGACCAATGCGCGGTGCG 600
QY 601 CGGATTCAGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CGGATTCAGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 661 CACCACTGCAACCGGAGCGTCTTGTGCGCTTCCGCGACCGCGACCGACCAAGGCGGAC 720
DB 661 CACCACTGCAACCGGAGCGTCTTGTGCGCTTCCGCGACCGCGACCGACCAAGGCGGAC 720
QY 721 ACCGATGCA 729
DB 721 ACCGATGCA 729

RESULT 2

AAV40170
ID AAV40170 standard; DNA, 1261 BP.

AAV40170;
AC

10-AUG-1999 (first entry)
DT

DNA sequence in Fig 59-60 of JP10155497.
DX

Carotenoid; pigment; canthaxanthin; R1534; crtB; prephycoene synthase;
KW crtI; phycoene desaturase; crtY; lycopene cyclase; crtW2396;
KW beta-carotene beta-oxygenase; food product; fermentation; ds.

Flavobacterium sp.
XX

JP10155497-A.
PN

16-JUN-1998.
PD

02-DEC-1997; 97JP-00348653.
PF

02-DEC-1996; 96EP-00810839.
PR

(HOFF) HOFFMANN LA ROCHE & CO AG F.
PA

WPI, 1998-391048/34.
DR

Preparation of carotenoid - comprises fermentation with transformed cell.
PT

Example 8; Fig 59-60; 80pp; Japanese.
PS

The invention describes the preparation of carotenoid pigments e.5.
XX canthaxanthin using a cell transformed by a vector having DNA sequences
XX (a) to (e) or substantially homologous sequences. (a) a DNA sequence
XX (crtE) coding GGP synthase of Flavobacterium sp. R1534; (b) a DNA
XX sequence (crtB) coding prephycoene synthase of Flavobacterium sp. R1534;
XX (c) a DNA sequence (crtI) coding phycoene desaturase of Flavobacterium
XX sp. R1534; (d) a DNA sequence (crtY) coding lycopene cyclase of
XX Flavobacterium sp. R1534, and (e) a DNA sequence (crtW2396) coding beta-
XX carotene beta-oxygenase of a microbe E-396 (FERM BP-4283). The carotenoid
XX or a carotenoid mixture can also be used in preparation of food products.
XX The method is an improved method of fermentation for carotenoid
XX production
XX

SQ Sequence 1261 BP; 188 A; 411 C; 406 G; 256 T; 0 U; 0 Other;
Best Local Similarity 100.0%; Score 729; DB 2; Length 1261;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGCATGCGCTTCCGCAAGCGAGATGACCGGCAACAGTTTGGCTCTCGGAGC 60
DB 40 ATGAGCGCATGCGCTTCCGCAAGCGAGATGACCGGCAACAGTTTGGCTCTCGGAGC 99
QY 61 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 120
DB 61 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 159
QY 100 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 159
DB 100 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 159
QY 121 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 180
DB 121 GGCATCATGCGCGCGTGGCGCTCCGTCAGTGCATGCGCTGGTTCTGACCGCGAGC 219
QY 181 TTCATCATGCGCGATGACGCGATGATGCGGTGCTGTCGCGCGCGCGCGCGCAT 240

```

Db      220 TTTCATATGCGCATGACCGGATGATGGTGTGTCGGGGGGCCCGCGGCAT 279
Qy      241 GCGGCGATGGGCGACGTTGCTCTGTGTATGCGGATTTTCTGCGGCAAGATGATC 300
Db      280 GCGGCGATGGGCGACGTTGCTCTGTGTATGCGGATTTTCTGCGGCAAGATGATC 339
Qy      301 GTCAAGACATGGGCGATGATGCGGATGCGGAAACGACGACCAATTTTCAGCAT 360
Db      340 GTCAAGACATGGGCGATGATGCGGATGCGGAAACGACGACCAATTTTCAGCAT 399
Qy      361 GCGGCGCGGTCGCTGTGTATGCGGCGGTCCTATTTTCGCTGGCGGAGGG 420
Db      400 GCGGCGCGGTCGCTGTGTATGCGGCGGTCCTATTTTCGCTGGCGGAGGG 459
Qy      421 CTGCTGCTGCGGTCATGCGGATGATGCGGTCCTATTTTCGCTGGAGTATC 480
Db      460 CTGCTGCTGCGGTCATGCGGATGATGCGGTCCTATTTTCGCTGGAGTATC 519
Qy      481 GTGCTCTTGTGCGGTCCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 540
Db      520 GTGCTCTTGTGCGGTCCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 579
Qy      541 TGCGTCGCGGACCGGCGGCGACGACGCGTCCTGCGGACCAATGCGGTCGTCG 600
Db      580 TGCGTCGCGGACCGGCGGCGACGACGCGTCCTGCGGACCAATGCGGTCGTCG 639
Qy      601 CGGATCAGCGACCGGTCGCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 660
Db      640 CGGATCAGCGACCGGTCGCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 699
Qy      661 CACCACTGACCGGACCGGTCCTGTGTATGCGGTCCTGCGGACCAATGCGGTCG 720
Db      700 CACCACTGACCGGACCGGTCCTGTGTATGCGGTCCTGCGGACCAATGCGGTCG 759
Qy      721 ACCGATGA 729
Db      760 ACCGATGA 768

RESULT 3
AAV40147
ID      AAV40147 standard; DNA; 729 BP.
XX      AAV40147;
XX      10-AUG-1999 (first entry)
XX      DNA sequence (crtW2396) encoding a beta-carotene beta-oxygenase.
XX      Carotenoid; pigment; canthaxanthin; R1534; crtB; prephytylene synthase;
XX      crtI; phytylene desaturase; crtY; lycopene cyclase; crtW2396;
XX      beta-carotene beta-oxygenase; food product; fermentation; ds.
XX      Flavobacterium sp.
OS      JP10155497-A.
XX      16-JUN-1998.
XX      02-DEC-1997; 97JP-00348653.
XX      02-DEC-1996; 96EP-00810839.
XX      (HOPF) HOPFMANN LA ROCHS & CO AG F.
XX      WPI; 1998-391048/34.
XX      P-RSDB; AAM69535.
XX      Preparation of carotenoid - comprises fermentation with transformed cell.
XX      Claim 1; Fig 61; 80pp; Japanese.
XX
```

```

CC      The invention describes the preparation of carotenoid pigments e.g.
CC      canthaxanthin using a cell transformed by a vector having DNA sequences
CC      (a) to (e) or substantially homologous sequences. (a) a DNA sequence
CC      (crtB) coding GPP synthase of Flavobacterium sp. R1534; (b) a DNA
CC      sequence (crtB) coding prephytylene synthase of Flavobacterium sp. R1534;
CC      (c) a DNA sequence (crtI) coding phytylene desaturase of Flavobacterium
CC      sp. R1534; (d) a DNA sequence (crtY) coding lycopene cyclase of
CC      Flavobacterium sp. R1534; and (e) a DNA sequence (crtW2396) coding beta-
CC      carotene beta-oxygenase of a microbe B-396 (FERM BP-4283). The carotenoid
CC      or a carotenoid mixture can also be used in preparation of food products.
CC      The method is an improved method of fermentation for carotenoid
CC      production
XX      SQ
XX      Sequence 729 BP; 100 A; 245 C; 231 G; 153 T; 0 U; 0 Other;
XX      Query Match 99.8%; Score 727.4; DB 2; Length 729;
XX      Best Local Similarity 99.9%; Pred. No. 3.7e-153;
XX      Matches 728; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 ATGAGGCGACATGCGGTCGCGGACGATGACCGGACCAAGTTTATGCTGCGG 60
Db      1 ATGAGGCGACATGCGGTCGCGGACGATGACCGGACCAAGTTTATGCTGCGG 60
Qy      61 GCGATCATGCGGCGGTCGCGGTCGCGGACGATGACCGGTCGCTGTGTATGAC 120
Db      61 GCGATCATGCGGCGGTCGCGGTCGCGGACGATGACCGGTCGCTGTGTATGAC 120
Qy      121 GCGATCATGCGGCGGTCGCGGTCGCGGACGATGACCGGTCGCTGTGTATGAC 180
Db      121 GCGATCATGCGGCGGTCGCGGTCGCGGACGATGACCGGTCGCTGTGTATGAC 180
Qy      181 TTTCATATGCGGATGACCGGATGACCGGTCGCTGTGTATGACCGGTCGCTGTGT 240
Db      181 TTTCATATGCGGATGACCGGATGACCGGTCGCTGTGTATGACCGGTCGCTGTGT 240
Qy      241 GCGGCGATGGGCGACGTTGCTGTGTATGCGGATTTTCTGCGGCAAGATGATC 300
Db      241 GCGGCGATGGGCGACGTTGCTGTGTATGCGGATTTTCTGCGGCAAGATGATC 300
Qy      301 GTCAAGACATGGGCGATGATGCGGATGCGGAAACGACGACCAATTTTCAGCAT 360
Db      301 GTCAAGACATGGGCGATGATGCGGATGCGGAAACGACGACCAATTTTCAGCAT 360
Qy      361 GCGGCGCGGTCGCTGTGTATGCGGCGGTCCTATTTTCGCTGGCGGAGGG 420
Db      361 GCGGCGCGGTCGCTGTGTATGCGGCGGTCCTATTTTCGCTGGCGGAGGG 420
Qy      421 CTGCTGCTGCGGTCATGCGGATGATGCGGTCCTATTTTCGCTGGAGTATC 480
Db      421 CTGCTGCTGCGGTCATGCGGATGATGCGGTCCTATTTTCGCTGGAGTATC 480
Qy      481 GTGCTCTTGTGCGGTCCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 540
Db      481 GTGCTCTTGTGCGGTCCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 540
Qy      541 TGCGTCGCGGACCGGCGGCGACGACGCGTCCTGCGGACCAATGCGGTCGTCG 600
Db      541 TGCGTCGCGGACCGGCGGCGACGACGCGTCCTGCGGACCAATGCGGTCGTCG 600
Qy      601 CGGATCAGCGACCGGTCGCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 660
Db      601 CGGATCAGCGACCGGTCGCTGTGTATGCGGTCCTATTTTCGCTGGAGTATC 660
Qy      661 CACCACTGACCGGACCGGTCCTGTGTATGCGGTCCTGCGGACCAATGCGGTCG 720
Db      661 CACCACTGACCGGACCGGTCCTGTGTATGCGGTCCTGCGGACCAATGCGGTCG 720
Qy      721 ACCGATGA 729
Db      721 ACCGATGA 729

RESULT 4
```

AD061160.
ID AD061160 standard; DNA; 729 BP.
XX
AC AD061160;
XX
DT 15-JUL-2004 (first entry)
XX
DE P. marcusii ketolase DNA.
XX
KM ds; ketocarotenoid; plant; ketolase; petal; flower-specific promoter;
KM petal-specific promoter; epsilon-cyclase; ec; chromoplast; hydroxylase;
KM beta-cyclase; bc; astaxanthin; canthaxanthin; echinenone; adonitrubin;
KM adonitrubin; ornamental; food; animal feed; supplement; pigment;
KM colouring; trout; salmon; shrimp; gene.
XX
OS Paracoccus marcusii.
XX
FH Key Location/Qualifiers
FT CDS 1..729
FT /tag= a
FT /product= "ketolase"
XX
PN DE10238978-A1.
XX
PD 04-MAR-2004.
XX
PF 20-AUG-2002; 2002DE-01038978.
XX
PR 20-AUG-2002; 2002DE-01038978.
XX
PA (SUNG-) SUNGENE GMBH & CO KGAA.
XX
DR MPI; 2004-215840/21.
XX
PT P-PSDB; AD061161.
XX
PT Method for preparing ketocarotenoids, useful e.g. as food or feed
XX supplements, by increasing, or introducing, ketolase activity in the
XX fruits of transgenic plants, also new nucleic acid constructs.
XX
PS Disclosure; SEQ ID NO 9; 77bp; German.
XX
XX This invention describes a novel method for preparing ketocarotenoids by
XX culturing genetically modified plants that, in comparison with the wild-
XX type, have altered ketolase activity in the fruits. The invention also
XX describes a nucleic acid construct comprising a fruit-specific promoter
XX linked functionally to a sequence that encodes the altered ketolase,
XX genetically modified plants in which the fruits have altered ketolase
XX activity and method for preparing such plants. The plants contain at
XX least one nucleic acid sequence that encodes the ketolase, especially
XX under control of a fruit-specific promoter, and particularly they contain
XX chromoplasts in the fruit. The ketolase-encoding nucleic acid is inserted
XX by standard methods, then the transgenic plants are cultivated, harvested
XX and ketocarotenoids are isolated from their fruits. Preferred carotenoids
XX include astaxanthin, canthaxanthin, echinenone (or its 3- or 3'-hydroxy
XX derivatives) adonitrubin or adonitrubin. The modified plants with
XX increased ketolase activity are used as ornamentals as food or animal
XX feed and for preparation of ketocarotenoids-containing extracts or for
XX preparing food/feed supplements, e.g., especially where the
XX ketocarotenoid is astaxanthin, as a pigment for colouring trout, salmon
XX and shrimp. The transgenic plants produce fruits with increased content
XX of ketocarotenoids. The invention describes a construct, vector pSKET02
XX which comprises, in pSUN5, a cassette containing the constitutive double
XX 35S cauliflower mosaic virus promoter; the rps chloroplast transit
XX peptide; the ketolase gene from Haemococcus pluvialis and a terminator.
XX It was used to transform tomato cells, using Agrobacterium tumefaciens,
XX and the infected cells regenerated to plants conventionally. One of the
XX resulting transgenic lines, CS13-24, produced fruits that contained
XX lycopene, beta-carotene, canthaxanthin, adonitrubin or astaxanthin, but
XX the last three were absent from wild-type fruits (which additionally
XX contained lutein, not detected in transgenic fruits).
XX
XX Sequence 729 BP; 103 A; 244 C; 231 G; 151 T; 0 U; 0 Other;

Query Match 95.8%; Score 698.6; DB 12; Length 729;
Best Local Similarity 97.4%; Pred. No. 1e-146;
Matches 710; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 ATGAGCGACATGCGCCCTGCGGAGGACATGACCGGACCAAGTTGATGTCGCGGAC 60
DB 1 ATGAGCGACATGCGCCCTGCGGAGGACATGACCGGACCAAGCTGATGTCGCGGAC 60
QY 61 GGCATCATGCGCGCGTGGCTGAGCTGATGATGCGCTGTGGTTCTGAGACGCGG 120
DB 61 GGCATCATGCGCGCGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 120
QY 121 GCGCATCCATCTGCGCGCTGCGGAAATTTCTGCGGCTGACCTGCTGCTGCTGCTG 180
DB 121 GCGCATCCATCTGCGCGCTGCGGAAATTTCTGCGGCTGACCTGCTGCTGCTGCTG 180
QY 181 TTGATCATGCGCGATGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 TTGATCATGCGCGATGAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 GCGGCGATGCGCGAGCTTGTCTGTGCTGTGATGCGGATTTTCTGCGCGAGATGATC 300
DB 241 GCGGCGATGCGCGAGCTTGTCTGTGCTGTGATGCGGATTTTCTGCGCGAGATGATC 300
QY 301 GTCAAGCATGAGCGCATCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCAT 360
DB 301 GTCAAGCATGAGCGCATCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCAT 360
QY 361 GCGGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GCGGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 CTGCTGCTGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 CTGCTGCTGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 TGGCTGCGCGACCG 600
DB 541 TGGCTGCGCGACCG 600
QY 601 CGGATCAGCGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CGGATCAGCGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CACGACCTGACCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 CACGACCTGACCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 ACCGATGA 729
DB 721 ACCGATGA 729
RESULT 5
AD061077
ID AD061077 standard; DNA; 729 BP.
XX
AC AD061077;
XX
DT 15-JUL-2004 (first entry)
XX
DE P. marcusii ketolase DNA.
XX
KM ds; ketocarotenoid; plant; ketolase; petal; flower-specific promoter;
KM petal-specific promoter; epsilon-cyclase; ec; chromoplast; hydroxylase;
KM beta-cyclase; bc; astaxanthin; canthaxanthin; echinenone; adonitrubin;
KM adonitrubin; ornamental; food; animal feed; supplement; pigment;
KM colouring; trout; salmon; shrimp; gene.
XX

OS Paracoccus marcusii.
 XX Key Location/Qualifiers
 FH CDS 1..729
 FT /*tag= a
 PN DE10238980-A1.
 XX
 PD 04-MAR-2004.
 PF 20-AUG-2002; 2002DE-01038980.
 XX
 PR 20-AUG-2002; 2002DE-01038980.
 XX
 PA (SUNG-) SUNGENE GMBH & CO KGAA.
 XX WPI; 2004-215842/21.
 DR P-PSDB; ADO61078.
 PT Method for preparing ketocarotenoids, useful e.g. as food or feed
 PT supplements, by increasing, or introducing, ketolase activity in the
 PT petals of transgenic plants, also new nucleic acid constructs.
 XX
 PS Disclosure; SEQ ID NO 9; 140pp; German.
 XX
 CC This invention describes a novel method for preparing ketocarotenoids by
 CC culturing genetically modified plants that, in comparison with the wild-
 CC type, have altered ketolase activity in the petals. The invention also
 CC describes a nucleic acid construct containing a nucleic acid that encodes
 CC ketolase, linked functionally to a flower- or petal-specific promoter,
 CC double-stranded RNA that comprises a sense strand, including a sequence
 CC that is essentially identical with at least part of the RNA epsilon-
 CC cyclase (ec) transcript or the promoter region of the ec gene and an
 CC antisense strand that is essentially complementary to the sense strand,
 CC an transgenic expression cassette comprising a plant-functional promoter
 CC linked to a nucleic acid that transcribes double-stranded RNA,
 CC genetically altered plant in which activity of ketolase in the petals is
 CC increased, if already present in the wild-type or introduced if absent
 CC from the wild type, genetically altered plant that has chromoplasts in
 CC the petals and contains at least one transgenic nucleic acid that encodes
 CC a ketolase. Particularly plants also have increased activity, relative to
 CC wild type, of hydroxylase and/or beta-cyclase (bc) activity, especially
 CC as a result of introducing the appropriate nucleic acid, and plants are
 CC also selected for highest activity of these two enzymes. The transgenic
 CC plants are cultivated, harvested and ketocarotenoids isolated from their
 CC petals. Preferred carotenoids include astaxanthin, canthaxanthin,
 CC echinone (or its 3- or 3'-hydroxy derivatives), adonixanthin or
 CC adonixanthin. The modified plants with increased ketolase activity are
 CC used as ornamentals, as food or animal feed and for preparation of
 CC ketocarotenoid-containing extracts or for preparing food/feed
 CC supplements, e.g., especially where the ketocarotenoid is astaxanthin, as
 CC a pigment for colouring trout, salmon and shrimps. The transgenic plants
 CC have increased content of ketocarotenoids. Vector p53KETO2 comprises, in
 CC PSUNs, a cassette containing the constitutive double 35S cauliflower
 CC mosaic virus promoter, the rbcS chloroplast transit peptide, the ketolase
 CC gene from Haemococcus pluvialis and a terminator. The vector was used to
 CC transform tomato cells, using Agrobacterium tumefaciens and the infected
 CC cells regenerated to plants conventionally. One of the resulting
 CC transgenic lines, CS13-8, produced orange flowers (yellow in the wild
 CC type) and its petals contained astaxanthin and adonixanthin, both absent
 CC from the wild type.
 XX
 XX Sequence 729 BP; 103 A; 244 C; 231 G; 151 T; 0 U; 0 Other;
 SQ
 Query Match 95.8%; Score 698.6; DB 12; Length 729;
 Best Local Similarity 97.4%; Pred. No. 1e-146;
 Matches 720; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 QY 1 ATGAGCGCACATCCCTGCCAAGGACGATCTGACCGCCACCACTTTGATCGTCTGGGC 60
 DB 1 ATGAGCGCACATCCCTGCCAAGGACGATCTGACCGCCACCACTTTGATCGTCTGGGC 60
 QY 61 GGATCATGCGCGCGTGGCGCTGATGTGATGCGCGTGTCTTGACGCGGCG 120

DB 61 GGATCATGCGCGCGTGGCGCTGATGTGATGCGCGTGTCTTGACGCGGCG 120
 QY 121 GCGATCCCATCTCTGCGCGGTCCGGAATTTCTCGGGCTCACTTGTGTGCTGGTCTG 180
 DB 121 GCGATCCCATCTCTGCGCGGTCCGGAATTTCTCGGGCTCACTTGTGTGCTGGTCTG 180
 QY 181 TTCATCATGCGGACATGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 240
 DB 181 TTCATCATGCGGACATGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 240
 QY 241 GCGGCGATGCGGACATGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 300
 DB 241 GCGGCGATGCGGACATGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 300
 QY 301 GTCAAGCACATGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 360
 DB 301 GTCAAGCACATGAGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 360
 QY 361 GCGGCGCGGATGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 420
 DB 361 GCGGCGCGGATGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 420
 QY 421 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 421 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 481 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 541 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 DB 541 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 601 CGATGAGGACCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 DB 601 CGATGAGGACCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 661 CACCACTGACCGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 720
 DB 661 CACCACTGACCGCGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 720
 QY 721 ACCGATGA 729
 DB 721 ACCGATGA 729
 RESULT 6
 AD038246
 ID AD038246 standard; DNA; 729 BP.
 AC AD038246;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 XX Paracoccus marcusii ketolase DNA.
 XX
 XX astaxanthin; plant; animal feed; fodder; fish; crustacean; bird;
 KW salmonid; shrimp; crab; hen; duck; geese; flamingo; ketolase;
 KW beta-carotene; canthaxanthin; hydroxylase; beta-cyclase; pigmentation;
 KW trout; salmon; shrimp; ds; gene.
 XX
 OS Paracoccus marcusii.
 XX
 XX Key Location/Qualifiers
 FH CDS 1..729
 FT /*tag= a
 FT /product= "ketolase"
 PN DE10258971-A1.
 XX
 PD 01-JUL-2004.

XX 16-DEC-2002; 2002DE-01058971.
 PF 16-DEC-2002; 2002DE-01058971.
 XX
 PR (SUNG-) SUNGENE GMBH & CO KGAA.
 XX
 PA WPI; 2004-554601/54.
 XX P-PSDB; ADQ38247.
 DR
 XX
 PT Use of astaxanthin-containing plant material, or extracts, from Tagetes
 PT for oral administration to animals, particularly for pigmentation of
 PT fish, crustacea, birds and their products.
 PS
 XX Disclosure; SEQ ID NO 9; 145bp; German.
 CC This invention describes the novel use of astaxanthin-containing plants
 CC (particularly the flower heads or petals) of the genus Tagetes or their
 CC parts for oral administration to animals. Methods are also described for
 CC a) the preparation of an animal feed composition by mixing standard
 CC fodder ingredients with astaxanthin-containing plants, b) for
 CC pigmentation of animals, or their products, by oral administration of
 CC astaxanthin-containing plants and c) animal feed composition or
 CC pigmentation agent that contains astaxanthin-containing plants. The
 CC plants, or their parts or extracts, are administered directly to animals,
 CC optionally after intermediate processing, or they are formulated with
 CC fodder components. Particularly they are administered to fish,
 CC crustaceans or birds, specifically salmonids, shrimps, crabs, hens,
 CC ducks, geese and flamingoes. The plants of the invention are preferably
 CC modified to provide ketolase activity, for conversion of beta-carotene to
 CC canthaxanthin, particularly with the highest level of expression in the
 CC flowers, e.g. by using a flower-specific promoter. The plants may also
 CC have increased activity of hydroxylase (for conversion of canthaxanthin
 CC to astaxanthin) and/or of beta-cyclase (to increase production of beta-
 CC carotene from gamma-carotene). Astaxanthin-containing compositions are
 CC used particularly for pigmentation of animals, preferably fish, crustacea
 CC and birds, or their products (meat, skin, feathers and eggs), most
 CC particularly trout, salmon and shrimp. Genetically modified Tagetes
 CC produce larger amounts of astaxanthin than the current source, Adonis
 CC aestivalis, and more cheaply. The invention uses expression vector
 CC pSKR202 which includes a cassette consisting of the double 35S promoter,
 CC the sequence for the pea rbc transit peptide, the sequence encoding a
 CC ketolase (beta-carotene-4-oxygenase) from Haemococcus pluvialis and the
 CC polyadenylation signal from cauliflower mosaic virus. It was used to
 CC transform leaves of Tagetes, by Agrobacterium-mediated transfer, then
 CC these regenerated to plants.
 CC
 XX
 XX
 SQ Sequence 729 BP; 103 A; 244 C; 231 G; 151 T; 0 U; 0 Other;
 Query Match 95.8%; Score 698.6; DB 12; Length 729;
 Best Local Similarity 97.4%; Pred. No. 1e-146;
 Matches 710; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 301 GTCAAGCACATGAGCCCATCATGCGCATGCGGAAACCGACGACGACCCAGATTGACCAT 360
 DB 301 GTCAAGCACATGAGCCCATCATGCGCATGCGGAAACCGACGACGACCCAGATTGACCAT 360
 QY 361 GCGGCGCGGTCCTGCTGTAGCCCGCTTCAATCGGCACCTTATTTGCGCTGGCGGAGGGG 420
 DB 361 GCGGCGCGGTCCTGCTGTAGCCCGCTTCAATCGGCACCTTATTTGCGCTGGCGGAGGGG 420
 QY 421 CTGCTGCTGCGCGTCATGTCGACGGTCTATGCGCTGATGTGTGGAGGATCGCTGATGTAC 480
 DB 421 CTGCTGCTGCGCGTCATGTCGACGGTCTATGCGCTGATGTGTGGAGGATCGCTGATGTAC 480
 QY 481 GTGCTCTTCTGACCGCTTCCCTGATCTGACGCTGATCTGACGCTGCTGCTGCTGCTGCT 540
 DB 481 GTGCTCTTCTGACCGCTTCCCTGATCTGACGCTGATCTGACGCTGCTGCTGCTGCTGCT 540
 QY 541 TGGCTGCGGACCG 600
 DB 541 TGGCTGCGGACCG 600
 QY 601 CGGATCAGCAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 DB 601 CGGATCAGCAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 661 CACCACTTGACACCGCGAGCGTGTCTTGTGCGCGCTGCGCAGCAGCAGCAGCAGCAGCAG 720
 DB 661 CACCACTTGACACCGCGAGCGTGTCTTGTGCGCGCTGCGCAGCAGCAGCAGCAGCAGCAG 720
 QY 721 ACCGATGTA 729
 DB 721 ACCGATGTA 729
 RESULT 7
 ADR03861
 ID ADR03861 standard; DNA; 729 BP.
 AC ADR03861;
 XX
 XX 21-OCT-2004 (first entry)
 DT
 XX Paracoccus marcusii ketolase coding sequence SEQ ID NO: 19.
 DE
 XX carotenoid; genetically modified; Blakeslea; nutrient; pigmentation;
 KW antioxidant; transgenic; ds; gene; ketolase; enzyme.
 XX
 XX Paracoccus marcusii.
 OS Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT 1..729
 FT CDS /tag= a
 FT /product= "ketolase"
 FT
 XX WO2004063359-A2.
 PN 29-JUL-2004.
 PD
 XX 09-JAN-2004; 2004WO-EP000099.
 PF
 XX 09-JAN-2003; 2003DE-01000649.
 PR 08-SEP-2003; 2003DE-01041271.
 XX
 XX (BADI) BASF AG.
 PA
 XX Matuschek M, Klein D, Heinekamp T, Schmidt A, Brakhage A;
 PI Achatz B;
 XX
 XX WPI; 2004-544088/52.
 DR P-PSDB; ADR03862.
 XX
 XX Preparing carotenoids or their precursors useful e.g. in cosmetics,
 PT pharmaceuticals, foods and animal feeds, comprises culturing genetically

PT modified Blakeslea.

XX Disclosure; SEQ ID NO 19; 486bp; German.

XX The present invention relates to a method of preparing carotenoids or
CC their precursors using genetically modified organisms of the genus
CC Blakeslea. The method is used for production of carotenoids, particularly
CC carotenes and xanthophylls, useful as animal and human nutrients, or
CC supplements, cosmetics, dermatological agents and pharmaceuticals,
CC particularly for pigmentation or colouring beverages, but also as
CC antioxidants. The present sequence is a polynucleotide used in the
CC exemplification of the invention.

XX Sequence 729 BP; 103 A; 244 C; 231 G; 151 T; 0 U; 0 Other;

Query Match 95.8%; Score 698.6; DB 13; Length 729;

Best Local Similarity 97.4%; Pred. No. 1e-146; Mismatches 19; Indels 0; Gaps 0;

Matches 710; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ATGAGGCGACATGCGCTGCGCCAAAGGAGATCTGACCGCCACCGATTGATCGTGGGGC 60
DB 1 ATGAGGCGACATGCGCTGCGCCAAAGGAGATCTGACCGCCACCGATTGATCGTGGGGC 60
QY 61 GGCATCATGCGCGCGCTGCGCTGCGCTGCGATGTCATGCGCTGCTTCTGACGCGGCG 120
DB 61 GGCATCATGCGCGCGCTGCGCTGCGCTGCGATGTCATGCGCTGCTTCTGACGCGGCG 120
QY 121 GGCATCATGCGCGCGCTGCGCGAATTTCTGCGGCTGACCTGCTGTGCGTGGCTG 180
DB 121 GGCATCATGCGCGCGCTGCGCGAATTTCTGCGGCTGACCTGCTGTGCGTGGCTG 180
QY 181 TTTCATCATGCGCGATGACCGGATGCGATGCGTGTGCTGCGCGCGCGCGCAT 240
DB 181 TTTCATCATGCGCGATGACCGGATGCGATGCGTGTGCTGCGCGCGCGCGCAT 240
QY 241 GCGGCGATGCGCGCGCTGCTGCTGCTGCTGATGCGGATTTCTGCGCGCAAGATGTC 300
DB 241 GCGGCGATGCGCGCGCTGCTGCTGCTGCTGATGCGGATTTCTGCGCGCAAGATGTC 300
QY 301 GTCAAGCATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GTCAAGCATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 CTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 CTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 TGGCTGCG 600
DB 541 TGGCTGCG 600
QY 601 CGGATCAGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CGGATCAGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CACCATCTGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 CACCATCTGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 ACCGCGATGA 729
DB 721 ACCGCGATGA 729

RESULT 8

ADRO3941
ID ADR03941 standard; DNA; 729 BP.

XX ADR03941;

XX 21-OCT-2004 (first entry)

XX P marcusii ketolase coding sequence.

XX gene: genetically modified; vector: carotenoid production; pigmentation;
KW nutrient; antioxidant; ds; ketolase; enzyme.

XX Paracoccus marcusii.
OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..729
FT /tag= a
FT /product= "ketolase"

PN MO2004063358-A1.

PD 29-JUL-2004.

PF 09-JAN-2004; 2004WO-EP000100.

PR 09-JAN-2003; 2003DE-01000649.

PR 08-SEP-2003; 2003DE-01041272.

PA (BAD1) BASF AG.

PI Matuschek M, Heinkeamp T, Schmidt A, Brakhage A;

DR WPI; 2004-544087/52.

DR P-PSDB; ADR03942.

PT Preparing genetically modified Blakeslea, useful for preparation of
PT carotenoids, useful as food additives, cosmetics or pharmaceuticals,
PT comprises transformation, optional homokaryotizing, and selection.

PS Disclosure; SEQ ID NO 19; 459bp; German.

XX The present invention relates to a method of preparing a genetically
CC modified organism of the genus Blakeslea, which comprises first
CC transforming at least one cell then optionally homokaryotizing the cells
CC so that cells are produced in which the nuclei are all simultaneously
CC altered in one or more genetic characteristics and these alterations are
CC expressed, and finally selection and culture of the modified cell(s). The
CC genetically modified organisms are used for the production of
CC carotenoids, particularly carotenes and xanthophylls, useful as animal
CC and human nutrients, or supplements, cosmetics and pharmaceuticals,
CC particularly for pigmentation or colouring beverages, but also as
CC antioxidants. The present sequence is a coding sequence shown in the
CC exemplification of the invention.

XX Sequence 729 BP; 103 A; 244 C; 231 G; 151 T; 0 U; 0 Other;

Query Match 95.8%; Score 698.6; DB 13; Length 729;

Best Local Similarity 97.4%; Pred. No. 1e-146; Mismatches 19; Indels 0; Gaps 0;

Matches 710; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ATGAGGCGACATGCGCTGCGCCAAAGGAGATCTGACCGCCACCGATTGATCGTGGGGC 60
DB 1 ATGAGGCGACATGCGCTGCGCCAAAGGAGATCTGACCGCCACCGATTGATCGTGGGGC 60
QY 61 GGCATCATGCGCGCGCTGCGCTGCGCTGCGATGTCATGCGCTGCTTCTGACGCGGCG 120
DB 61 GGCATCATGCGCGCGCTGCGCTGCGCTGCGATGTCATGCGCTGCTTCTGACGCGGCG 120
QY 121 GGCATCATGCGCGCGCTGCGCGAATTTCTGCGGCTGACCTGCTGTGCGTGGCTG 180
DB 121 GGCATCATGCGCGCGCTGCGCGAATTTCTGCGGCTGACCTGCTGTGCGTGGCTG 180

QY	421	CTGCTGCTGACCCGCTATCTGTAACGCTCTATGCGCTGATGTTGGAGGATGCTGATGCTAC	480
Db	421	CTGCTGCTGACCCGCTATCTGTAACGCTCTATGCGCTGATGTTGGAGGATGCTGATGCTAC	480
QY	481	GTGCTCTTCTGGCCGCTTGGCCGCTGATCTTGCCTGATCCAGCTGTTCTGTTCCGATC	540
Db	481	GTGCTCTTCTGGCCGCTTGGCCGCTGATCTTGCCTGATCCAGCTGTTCTGTTCCGATC	540
QY	541	TGGCTGCGGCACATCCGCTCCGGCCACACACGCGTTTCCGGACCGACAAATGCGCGTGTCTG	600
Db	541	TGGCTGCGGCACATCCGCTCCGGCCACACACGCGTTTCCGGACCGACAAATGCGCGTGTCTG	600
QY	601	CGGATCAGGAGACCCCGCTGTGCGCTGACGACCTGCTTCACTTTGGCGGTTATCATCAGAA	660
Db	601	CGGATCAGGAGACCCCGCTGTGCGCTGACGACCTGCTTCACTTTGGCGGTTATCATCAGAA	660
QY	661	CACCACTTGACCCGACCGGCTGCTTGTGTGTGCGCTGCCGACACCCGACCAAGGAGGAC	720
Db	661	CACCACTTGACCCGACCGGCTGCTTGTGTGTGCGCTGCCGACACCCGACCAAGGAGGAC	720
QY	721	ACCGCATGA 729	
Db	721	ACCGCATGA 729	
RESULT 10			
ID	AD061073	standard; DNA; 729 BP.	
AC	AD061073;		
XX			
DT	15-JUL-2004	(first entry)	
XX			
DE	A. aurantiacum ketolase DNA.		
XX			
KW	ds; ketocarotenoid; plant; ketolase; petal; flower-specific promoter;		
KW	petal-specific promoter; epsilon-cyclase; ec; chromoplast; hydroxylase;		
KW	beta-cyclase; bc; aetaxanthin; canthaxanthin; echinone; adonirubin;		
KW	adonixanthin; ornamental; food; animal feed; supplement; pigment;		
KW	colouring; trout; salmon; shrimp; gene.		
XX			
OS	Agrobacterium aurantiacum.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	1..729	
FT-		/*tag= a	
PN	DE10238980-A1.		
XX			
PD	04-MAR-2004.		
XX			
PF	20-AUG-2002; 2002DE-01038980.		
XX			
PPR	20-AUG-2002; 2002DE-01038980.		
XX			
PA	(SUNG-) SUNGENE GMBH & CO KGAA.		
XX			
DR	WPI; 2004-215842/21.		
XX	P-PSDB; AD061074.		
PT	Method for preparing ketocarotenoids, useful e.g. as food or feed		
PT	supplements, by increasing, or introducing, ketolase activity in the		
PT	petals of transgenic plants, also new nucleic acid constructs.		
XX			
PS	Disclosure; SEQ ID NO 5; 140bp; German.		
CC	This invention describes a novel method for preparing ketocarotenoids by		
CC	culturing genetically modified plants that, in comparison with the wild-		
CC	type, have altered ketolase activity in the petals. The invention also		
CC	describes a nucleic acid construct containing a nucleic acid that encodes		
CC	ketolase, linked functionally to a flower- or petal-specific promoter,		
CC	double-stranded RNA that comprises a sense strand, including a sequence		

that is essentially identical with at least part of the RNA epsilon-cyclase (ec) transcript or the promoter region of the ec gene and an antisense strand that is essentially complementary to the sense strand, an transgenic expression cassette comprising a plant-functional promoter linked to a nucleic acid that transcribes double-stranded RNA, CC genetically altered plant in which activity of ketolase in the petals is increased, if already present in the wild-type or introduced if absent from the wild type, genetically altered plant that has chromoplasts in the petals and contains at least one transgenic nucleic acid that encodes a ketolase. Particularly plants also have increased activity, relative to wild type, of hydroxylase and/or beta-cyclase (bc) activity, especially CC as a result of introducing the appropriate nucleic acid, and plants are also selected for highest activity of these two enzymes. The transgenic CC plants are cultivated, harvested and ketocarotenoids isolated from their petals. Preferred carotenoids include astraanthin, cantaxanthin, CC echinone (or its 3- or 3'-hydroxy derivatives), adonirubin or CC adonixanthin. The modified plants with increased ketolase activity are used as ornamentals, as food or animal feed and for preparation of CC ketocarotenoids-containing extracts or for preparing food/feed supplements, e.g., especially where the ketocarotenoid is astraanthin, as CC a pigment for colouring trout, salmon and shrimps. The transgenic plants have increased content of ketocarotenoid. Vector pSKRTO comprises, in CC pSUNS, a cassette containing the constitutive double 35S cauliflower mosaic virus promoter, the rbs chloroplast transit peptide, the ketolase gene from Haemophilus pluvialis and a terminator. The vector was used to transform tomato cells, using Agrobacterium tumefaciens and the infected CC cells regenerated to plants conventionally. One of the resulting transgenic lines, Gs13-8, produced orange flowers (yellow in the wild CC type) and its petals contained astraanthin and adonixanthin, both absent from the wild type.

SQ Sequence 729 BP; 103 A; 249 C; 229 G; 148 T; 0 U; 0 Other;

Query Match	95.6%;	Score 697;	DB 12;	Length 729;
-------------	--------	------------	--------	-------------

Matches 709; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

OY	1	TTGAGCGGCAATGCCCTTCGACAGGCAATCTGACACGGCAACAGATTGATCGTCCGGG	60
Db	1	ATGAGCGGCAATGCCCTTCGACAGGCAATCTGACACGGCAACAGATTGATCGTCCGGG	60
OY	61	GGCATCATCGCGCGGTGGCTGGACCTGTGATGTCATGCGCTGTGGTTTCTGACACGGAG	120
Db	61	GGCATCATCGCGCGCTTGGCTGGACCTGTGATGTCATGCGCTGTGGTTTCTGACACGGAG	120
OY	121	GCGCATCCCATCTTGGCGCGTCCGCAATTTCTGAGGGCTGACCTGGCTGTCGGTCCGCTTG	180
Db	121	GCGCATCCCATCTTGGCGCATCGCAAAATTTCTCGAGGGCTGACCTGGCTGTCGGTCCGCTTG	180
OY	181	TTTCATCATCGCGCATGACGCGATGCGATGCGTCCGTGTCGCGGGCGCGCCGCGCCAAAT	240
Db	181	TTTCATCATCGCGCATGACGCGATGCGAGCGATGCGCGGGTCCGTGTCGCGGGCGCGCCAAAT	240
OY	241	GCGCGATGGGCGCACTTGTCTGTGGGCTGTATGCCGCAATTTTCTGGGGCAAGATGATC	300
Db	241	GCGCGATGGGCGCACTTGTCTGTGGGCTGTATGCCGCAATTTTCTGGGGCAAGATGATC	300
OY	301	GTCAAGCAGATGGGCCATCATCGCATGCGCGGAACCGAGCAGACCCAGATTTTCAGCAAT	360
Db	301	GTCAAGCAGATGGGCCATCATCGCATGCGCGGAACCGAGCAGACCCAGATTTTCAGCAAT	360
OY	361	GCGCGCCCGGTCCGCTGTAGCGCCCGCTTCATCGGACCTATTTCCGCTGGCGCGAGGGG	420
Db	361	GCGCGCCCGGTCCGCTGTAGCGCCCGCTTCATCGGACCTATTTCCGCTGGCGCGAGGGG	420
OY	421	CTGCTGCTGCCCGGTATGTGACGGTCTATAGCGCTGATGTTGGGGGATGCTGCGATGTAC	480
Db	421	CTGCTGCTGCCCGGTATGTGACGGTCTATAGCGCTGATGTTGGGGGATGCTGCGATGTAC	480
OY	481	GTGATCTTCTGGCGGTGCGGTGCATCTCGGGGTGATCCAGCTGTGGTGTTCGGCAATC	540
Db	481	GTGATCTTCTGGCGGTGCGGTGCATCTCGGGGTGATCCAGCTGTGGTGTTCGGCAATC	540

PI Matuschek M, Heinekamp T, Schmidt A, Brakhage A;
 XX WPI: 2004-54087/52.
 DR P-PSDB; ADR03938.
 XX
 PT Preparing genetically modified *Blakeslea*, useful for preparation of
 PT carotenoids, useful as food additives, cosmetics or pharmaceuticals,
 PS comprises transformation, optional homokaryotizing, and selection.
 XX
 PS Disclosure; SEQ ID NO 15; 459bp; German.
 XX
 CC The present invention relates to a method of preparing a genetically
 CC modified organism of the genus *Blakeslea*, which comprises first
 CC transforming at least one cell then optionally homokaryotizing the cells
 CC so that cells are produced in which the nuclei are all simultaneously
 CC altered in one or more genetic characteristics and these alterations are
 CC expressed, and finally selection and culture of the modified cell(s). The
 CC genetically modified organisms are used for the production of
 CC carotenoids, particularly carotenes and xanthophylls, useful as animal
 CC and human nutrients, or supplements, cosmetics and pharmaceuticals,
 CC particularly for pigmentation or colouring beverages, but also as
 CC antioxidants. The present sequence is a coding sequence shown in the
 CC exemplification of the invention.
 XX
 SQ Sequence 729 BP; 103 A; 249 C; 229 G; 148 T; 0 U; 0 Other;
 Query Match 95.6%; Score 697; DB 13; Length 729;
 Best Local Similarity 97.3%; Pred. No. 2,3e-146;
 Matches 709; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 1 ATGAGGCGACATGCCCCCTGCGCAAGGCGATGACCGCCACCAAGTTGATGCTCGGCG 60
 DB 1 ATGAGGCGACATGCCCCCTGCGCAAGGCGATGACCGCCACCAAGCTGATCTCGGCGC 60
 QY 61 GGCATCATGCGCGCGGCTGCGCTGCGATGTCGATGCGCTGCTGTTGAGACGCGG 120
 DB 61 GGCATCATGCGCGCTGCGCTGCGCTGCGATGTCGATGCTGCTGTTGAGACGCGG 120
 QY 121 GGCATATCCCATCTGCGCGGCTGCGCAATTTCTGCGGCGTACCTGCGTGCCTGCTG 180
 DB 121 GGCATATCCCATCTGCGCGGCTGCGCAATTTCTGCGGCGTACCTGCGTGCCTGCTG 180
 QY 181 TTTCATCATGCGCGATGACCGCGATGCGGCTGCTGCGCGCGCGCGCGCAT 240
 DB 181 TTTCATCATGCGCGATGACCGCGATGCGGCTGCTGCGCGCGCGCGCGCAT 240
 QY 241 GGGGCGATGCGCGACGCTTCTGCTGCGCTGCTGATGCGGATTTCTGCGGCGCAATGATC 300
 DB 241 GGGGCGATGCGCGACGCTTCTGCTGCGCTGCTGATGCGGATTTCTGCGGCGCAATGATC 300
 QY 301 GTCAAGCATGCGCGCATCATGCGCGCGGAAACCGACGACCGACCAATTTGACCAT 360
 DB 301 GTCAAGCATGCGCGCATCATGCGCGCGGAAACCGACGACCGACCAATTTGACCAT 360
 QY 361 GCGCGCGCGCTGCGCTGCTGTAACGCGCGCTTTCATGCGACCTATTTGCGTGCAGAGGG 420
 DB 361 GCGCGCGCGCTGCGCTGCTGTAACGCGCGCTTTCATGCGACCTATTTGCGTGCAGAGGG 420
 QY 421 CTGCTGCTGCGCGCTGCTGTAACGCTGTAATGCGGCTGATGTTGGGGGATGCTGATGTC 480
 DB 421 CTGCTGCTGCGCGCTGCTGTAACGCTGTAATGCGGCTGATGTTGGGGGATGCTGATGTC 480
 QY 481 GTGCTCTTGGCGCGTGTGCGTGTGATCTGCGCGTGCATCGAGTGTGCTGTTGCGGCATC 540
 DB 481 GTGCTCTTGGCGCGTGTGCGTGTGATCTGCGCGTGCATCGAGTGTGCTGTTGCGGCATC 540
 QY 541 TGGCTGCGCGACCGCGCGCGCGCGACGAGGCTTCCCGGACCGCGACCATGCGCGTGTGCG 600
 DB 541 TGGCTGCGCGACCGCGCGCGCGCGCGACGAGGCTTCCCGGACCGCGACCATGCGCGTGTGCG 600
 QY 601 CGGATCATGCGACCGCGCTGCTGCTGCTGATCCTGTTCACTTTGGCGGTTATCATCGAA 660
 DB 601 CGGATCATGCGACCGCGCTGCTGCTGCTGATCCTGTTCACTTTGGCGGTTATCATCGAA 660

QY 661 CACCACCTGCAACCGACCGGTGCTTGTGCGCTTCCCGACACCGCGACCAAGGGGAC 720
 DB 661 CACCACCTGCAACCGACCGGTGCTTGTGCGCTTCCCGACACCGCGACCAAGGGGAC 720
 QY 721 ACCGCATGA 729
 DB 721 ACCGCATGA 729
 RESULT 14
 AA099485
 ID AA099485 standard; DNA; 639 BP.
 AC AA099485;
 XX
 DT 27-AUG-2003 (revised)
 DT 28-FEB-1996 (first entry)
 XX
 DB 3 hydroxy-beta-ionone ring methylene to keto converting peptide DNA.
 XX
 KW Xanthophyll; astaxanthine; methylene; keto group; conversion;
 XX 3-hydroxy-beta-ionone ring, ds.
 OS Agrobacterium aurantiacum.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..639
 FT /tag= a
 FT /product= "beta-ionone_ring_4-methylene_to_4-keto
 FT _group_converting_polypeptide."
 XX
 PD WO9518220-A1.
 XX
 PF 06-JUL-1995.
 XX
 PF 26-DEC-1994; 94WO-JP002220.
 XX
 PR 27-DEC-1993; 93JP-00348737.
 PR 05-SEP-1994; 94JP-00235917.
 XX
 PA (KIRI) KIRIN BEER KK.
 PA (MARI-) MARINE BIOTECHNOLOGY INST CO LTD.
 PI Misawa N, Kondo K, Kajiwara S, Yokoyama A;
 XX
 DR WPI: 1995-246386/32.
 DR P-PSDB; AAR79058.
 PT DNA's encoding xanthophyll(s) - esp. asta:xanthin and other
 PT xanthophyll(s) using e.g. *E. coli*.
 XX
 PS Claim 2; Page 76-78; 131pp; Japanese.
 XX
 CC Q999485-AA099489 encode xanthophyll polypeptides. These polypeptides are
 CC capable of converting the 4-methylene group of a 3-hydroxy-beta-ionone
 CC ring to a 4-keto group in doing so these peptides also add a hydroxyl
 CC group to the 3-position carbon-atom of the 4-keto-beta-ionone ring. The
 CC DNA sequences may be used in the production of astaxanthine and other
 CC keto gp. contg. xanthophylls, the sequences may also be used to transform
 CC certain yeasts and other microorganisms. (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX
 SQ Sequence 639 BP; 88 A; 215 C; 204 G; 132 T; 0 U; 0 Other;
 Query Match 83.9%; Score 611.8; DB 2; Length 639;
 Best Local Similarity 97.3%; Pred. No. 2.6e-127;
 Matches 622; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 91 GTGCATGCGCGTGTGCTTCTGGAAGCGCGCGCATCCATCTGCGCGTGCAGATTTC 150
 DB 1 GTGCATGCGCGTGTGCTTCTGGAAGCGCGCGCATCCATCTGCGCGTGCAGATTTC 60

DR MPI; 1999-093030/09.
DR P-RSDB; AAM67882.

XX New carotenoid glucoside(s) - used as food additives.
XX
PS Diaclosure; Page 10-11; 26pp; Japanese.
XX

The present sequence encodes a protein involved in carotenoid biosynthesis. The specification describes astaxanthin diglucosides and adonixanthin-3'-glucosides. The specification also describes a method for the preparation of a carotenoid glycoside, in which all, or part of, carotenoid biosynthesis genes crtE, crtB, crtl, ctyI, ctyX or crtW are introduced to a microbe or plant and expressed. The transformed organism is cultured and astaxanthin diglucosides, adonixanthin-3'-glucosides, and/or astaxanthin monoglucoside are collected. The carotenoid glucosides are used as food additives

SQ Sequence 639 BP; 88 A; 215 C; 204 G; 132 T; 0 U; 0 Other;

Query Match 83.9%; Score 611.8; DB 2; Length 639;
Best Local Similarity 97.3%; Pred. No. 2,6e-127;
Matches 622; Conservative 0; Mismatches 17; Indels 0; Gaps 0

Dy 91 GTGCATGCGCTGTGGTTTCTTGACAGCGAGCGCATCCCATCTTCGCGCAATTTC 150
Dd 1 GTGCATGCGCTGTGGTTTCTTGACAGCGAGCGCATCCCATCTTCGCGCAATTTC 60

Dy 151 CTGAGGACTGAACCTGGCTGTGCTGGTCTGTTCATCATGCGCATAAGCGATGCATGGG 210
Dd 61 CTGAGGACTGAACCTGGCTGTGCTGGTCTGTTCATCATGCGCATAAGCGATGCATGGG 120

Dy 211 TCGGTGTGTGCCGGGGGCCGCCGCCGCACAATGCGCGATGGCCACGCTTGTCTTGGCTG 270
Dd 121 TCGGTGTGTGCCGGGGGCCGCCGCCGCACAATGCGCGATGGCCACGCTTGTCTTGGCTG 180

Dy 271 TAGCCGGAAATTTTCGGGCGCAAAGTAGTCGAAGAAGCATAGGCCAATATGCGCATGCC 330
Dd 181 TAGCCGGAAATTTTCGGGCGCAAAGTAGTCGAAGAAGCATAGGCCAATATGCGCATGCC 240

Dy 331 GGAACGAGACGACACCACNAATTTTCACACATGAGCGGCCCGATCGCTGGTAGCCCCGCTTC 390
Dd 241 GGAACGAGACGACACCACCCGATTTTCACACATGAGCGGCCCGATCGCTGGTAGCCCCGCTTC 300

Dy 391 ATGCGACACTATTTTCGGCTGTGGCGCGAAGGAGGCGTGTCTGTCCGTTCATGCGTCTAT 450
Dd 301 ATGCGACACTATTTTCGGCTGTGGCGCGAAGGAGGCGTGTCTGTCCGTTCATGCGTCTAT 360

Dy 451 GGCGCTATGTGGGGAGATGCGTGTGAATGTAACGTGCTTCTGAGCGGTGCGTGCATGCTG 510
Dd 361 GGCGCTATGTTGGGGAGATGCGTGTGAATGTAACGTGCTTCTGAGCGGTGCGTGCATGCTG 420

Dy 511 GCCTGCATCACGCTGTTGCTGTTGGGCAATCTTGCGCGACCGGCCCGGACGACGCG 570
Dd 421 GCCTGCATCACGCTGTTGCTGTTGGGCAATCTTGCGCGACCGGCCCGGACGACGCG 480

Dy 571 TTCCTCGAACCGCCACCAATGCGCGGTGTGTGCGGATCAGCGACCCCGTGTGTGTGAC 630
Dd 481 TTCCTCGAACCGCCACCAATGCGCGGTGTGTGCGGATCAGCGACCCCGTGTGTGAC 540

Dy 631 TGCTTTCACTTTGGCGGTTATATATCAAGAACCAACCTGACCCGAGGGTCTTGGTGG 690
Dd 541 TGCTTTCACTTTGGCGGTTATATATCAAGAACCAACCTGACCCGAGGGTCTTGGTGG 600

Dy 691 CGCCTGCCAGACCCGCGACCCAAGGGGAGACACCGCATGA 729
Dd 601 CGCCTGCCAGACCCGCGACCCAAGGGGAGACACCGCATGA 639

Search completed: June 21, 2005, 20:32:42
Job time : 479.072 secs

This Page Blank (uspto)


```
QY 361 GCGGCGCCGGTCCGCTGTGACGCCGCTTCATCGGCACTTATTCGGCTGGCGGAGG 420
|
|
|
DB 361 GCGGCGCCGGTCCGCTGTGACGCCGCTTCATCGGCACTTATTCGGCTGGCGGAGG 420
|
|
|
QY 421 CTGCTGTGCGCCGTCATCGTGAACGCTGTATGCGCTGTATGCGGAGTCTGTGATGAC 480
|
|
|
DB 421 CTGCTGTGCGCCGTCATCGTGAACGCTGTATGCGCTGTATGCGGAGTCTGTGATGAC 480
|
|
|
QY 481 GTGGCTCTTGTGCGCGTGTGCGATCCGTCGCGTGTGATCAGCTGTGCTGTGTCGACATC 540
|
|
|
DB 481 GTGGCTCTTGTGCGCGTGTGCGATCCGTCGCGTGTGATCAGCTGTGCTGTGTCGACATC 540
|
|
|
QY 541 TGGCTGCGCGACCGCCCGGCGACGAGCGGTTCCCGGACCGGCAACATCGCGGTCTGCG 600
|
|
|
DB 541 TGGCTGCGCGACCGCCCGGCGACGAGCGGTTCCCGGACCGGCAACATCGCGGTCTGCG 600
|
|
|
QY 601 CGGATCAGCGACCCCGTGTGCTGTGACCTGTGCTTCACTTTGGCGGTTATCATCAGAA 660
|
|
|
DB 601 CGGATCAGCGACCCCGTGTGCTGTGACCTGTGCTTCACTTTGGCGGTTATCATCAGAA 660
|
|
|
QY 661 CACCACTTGACACCCGAGCGCTGTGTGTGGCGCTGGCCAGACCCCGGCAAGGGGAGC 720
|
|
|
DB 661 CACCACTTGACACCCGAGCGGTGTGTGTGGCGCTGGCCAGACCCCGGCAAGGGGAGC 720
|
|
|
QY 721 ACCGCATGA 729
|
|
|
DB 721 ACCGCATGA 729
|
|
|
```

```
RESULT 2
US-08-980-832-30
; Sequence 30, Application US/08980832B
; Patent No. 6291204
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Teyganov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/08/980,832B
; CURRENT FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1261
; TYPE: DNA
; ORGANISM: Alcaligenes PC-1
US-08-980-832-30
```

```
Query Match
Best Local Similarity 100.0%; Score 729; DB 3; Length 1261;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAGCGACATGCGCTGTGCGCAAGGAGATCTGACCGGCAACGATTGATGCTTCGGGC 60
|
|
|
DB 40 ATGAGCGACATGCGCTGTGCGCAAGGAGATCTGACCGGCAACGATTGATGCTTCGGGC 99
|
|
|
QY 61 GGCATCATGCGCGCGTGTGCGCTGTGATGTCATGCGCTGTGCTTTTGTGACCGGGG 120
|
|
|
DB 100 GGCATCATGCGCGCGTGTGCGCTGTGATGTCATGCGCTGTGCTTTTGTGACCGGGG 159
|
|
|
QY 121 GGCATCATGCGCGCGTGTGCGCAATTTCTGCGGCTGACCTGTGCTGTGCTGTG 180
|
|
|
DB 160 GGCATCATGCGCGCGTGTGCGCAATTTCTGCGGCTGACCTGTGCTGTGCTGTG 219
|
|
|
QY 181 TTTCATCATGCGCGATGACCGGATGATGAGTGGTGGTGTGTCGCGGCGCGCGCAAT 240
|
|
|
DB 220 TTTCATCATGCGCGATGACCGGATGATGAGTGGTGGTGTGTCGCGGCGCGCGCAAT 279
|
|
|
QY 241 GGGGCGATGGGCGAGTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
|
|
|
DB 280 GGGGCGATGGGCGAGTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 339
|
|
|
```

```
QY 301 GTCAAGCATATGAGCCCATCATCGCCATGCGGAAACGACGACGACCCAGATTTTGCACAT 360
|
|
|
DB 340 GTCAAGCATATGAGCCCATCATCGCCATGCGGAAACGACGACGACCCAGATTTTGCACAT 399
|
|
|
QY 361 GCGGCGCCGGTCCGCTGTGTAAGCCCGCTTCATTCGGACCTTATTTGGCTGGCGGAGG 420
|
|
|
DB 400 GCGGCGCCGGTCCGCTGTGTAAGCCCGCTTCATTCGGACCTTATTTGGCTGGCGGAGG 459
|
|
|
QY 421 CTGCTGTGCGCCGTCATCGTGAACGCTGTATGCGCTGTATGCGGAGTCTGTGATGAC 480
|
|
|
DB 460 CTGCTGTGCGCCGTCATCGTGAACGCTGTATGCGCTGTATGCGGAGTCTGTGATGAC 519
|
|
|
QY 481 GTGGCTCTTGTGCGCGTGTGCGATCCGTCGCGTGTGATCAGCTGTGCTGTGTCGACATC 540
|
|
|
DB 520 GTGGCTCTTGTGCGCGTGTGCGATCCGTCGCGTGTGATCAGCTGTGCTGTGTCGACATC 579
|
|
|
QY 541 TGGCTGCGCGACCGCCCGGCGACGAGCGGTTCCCGGACCGGCAACATGCGCGGTCTGCG 600
|
|
|
DB 580 TGGCTGCGCGACCGCCCGGCGACGAGCGGTTCCCGGACCGGCAACATGCGCGGTCTGCG 639
|
|
|
QY 601 CGGATCAGCGACCCCGTGTGCTGTGACCTGTGCTTCACTTTGGCGGTTATCATCAGAA 660
|
|
|
DB 640 CGGATCAGCGACCCCGTGTGCTGTGACCTGTGCTTCACTTTGGCGGTTATCATCAGAA 699
|
|
|
QY 661 CACCACTTGACACCCGAGCGGTGTGTGTGGCGCTGGCCAGACCCCGGCAAGGGGAGC 720
|
|
|
DB 700 CACCACTTGACACCCGAGCGGTGTGTGTGGCGCTGGCCAGACCCCGGCAAGGGGAGC 759
|
|
|
QY 721 ACCGCATGA 729
|
|
|
DB 760 ACCGCATGA 768
|
|
|
```

```
RESULT 3
US-09-920-923B-30
; Sequence 30, Application US/09920923B
; Patent No. 6677134
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Teyganov, Yuri
; TITLE OF INVENTION: Fermentative Carotenoid Production
; FILE REFERENCE: 15464 US (C38435/125944)
; CURRENT APPLICATION NUMBER: US/09/920,923B
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 1261
; TYPE: DNA
; ORGANISM: Alcaligenes PC-1
US-09-920-923B-30
```

```
Query Match
Best Local Similarity 100.0%; Score 729; DB 4; Length 1261;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAGCGACATGCGCTGTGCGCAAGGAGATCTGACCGGCAACGATTGATGCTTCGGGC 60
|
|
|
DB 40 ATGAGCGACATGCGCTGTGCGCAAGGAGATCTGACCGGCAACGATTGATGCTTCGGGC 99
|
|
|
QY 61 GGCATCATGCGCGCGTGTGCGCTGTGATGTCATGCGCTGTGCTTTTGTGACCGGGG 120
|
|
|
DB 100 GGCATCATGCGCGCGTGTGCGCTGTGATGTCATGCGCTGTGCTTTTGTGACCGGGG 159
|
|
|
QY 121 GGCATCATGCGCGCGTGTGCGCAATTTCTGCGGCTGACCTGTGCTGTGCTGTG 180
|
|
|
DB 160 GGCATCATGCGCGCGTGTGCGCAATTTCTGCGGCTGACCTGTGCTGTGCTGTG 219
|
|
|
QY 181 TTTCATCATGCGCGATGACCGGATGATGAGTGGTGGTGTGTCGCGGCGCGCGCAAT 240
|
|
|
DB 220 TTTCATCATGCGCGATGACCGGATGATGAGTGGTGGTGTGTCGCGGCGCGCGCAAT 279
|
|
|
```


APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2886 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-335-919-7

Query Match 95.6%; Score 697; DB 3; Length 2886;
Best Local Similarity 97.3%; Pred. No. 6,4e-159;
Matches 709; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

1 ATGAGCGCACATCCCTGCGCCAGAGGATGTGACCGCCACAGTTTGATGCTCGGGC 60
139 ATGAGCGCACATCCCTGCGCCAGAGGATGTGACCGCCACAGTTTGATGCTCGGGC 198
61 GGCATATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 120
199 GGCATATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 258
121 GGCATATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 180
259 GGCATATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 318
181 TTCATATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 240
319 TTCATATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 378
241 GCGCGATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 300
379 GCGCGATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 438
301 GTCACACATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 360
439 GTCACACATGCGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 498
361 GCGCGCGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 420
499 GCGCGCGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 558
421 CTGCTGCTGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 480
559 CTGCTGCTGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 618
481 GTGCTCTTCTGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 540
619 GTGCTCTTCTGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 678
541 TGGCTGCGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 600
679 TGGCTGCGCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 738
601 CGGATCAGCGACCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 660
739 CGGATCAGCGACCGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 798

661 CACCACTGACCGGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 720
799 CACCACTGACCGGCGTGGCTGGCGCGTGGCTGGCGCTGATGTCATGTCATGCTGCTTTCGACGCGCG 858

721 ACCGATGA 729
859 ACCGATGA 867

RESULT 7
US-08-663-310-1
Sequence 1, Application US/08663310
Patent No. 5811273
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5811273hiko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIMAWA, Sueumu
TITLE OF INVENTION: YOKOYAMA, Akihiro
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSER: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 639 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..636
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..636
US-08-663-310-1

Query Match 83.9%; Score 611.8; DB 1; Length 639;
Best Local Similarity 97.3%; Pred. No. 1.6e-138;
Matches 62; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy	9	GTGATGACCGGTGAGTTTCTTGACACGGGAGGGACATCCCATCTGGAGGTGGCAATTTC	150
Db	1	GTGATGACCGGTGAGTTTCTTGACACGGGAGGGACATCCCATCTGGAGGTGGCAAAATTC	60
Qy	151	CTGGAGCTGACCTGAGCTGTGCGTCTGTTTCATCATTCGCGATGACGCGATGCATGAG	210
Db	61	CTGGAGGTGACCTGAGCTGTGCGTCTGTTTCATCATTCGCGATGACGCGATGCATGAG	120
Qy	211	TTCGTCGTGCCCCGGGGGGCCCCGGCGCCCATATGGCGGATATGGGCGACCTTGTCTTGACTG	270
Db	121	TTCGTCGTGCCCCGGGGGGCCCCGGCGCCCATATGGCGGATATGGGCGACCTTGTCTTGACTG	180
Qy	271	TATGCCGGAATTTTCTGGGCGAAAGATGATGCTCAAGACATATGAGCCATATCGCATGACC	330
Db	181	TATGCCGGAATTTTCTGGGCGAAAGATGATGCTCAAGACATATGAGCCATATCGCATGACC	240
Qy	331	GGAACCGACGACGACCCGACATTTTCACATATGGCGGCCGCGTCCGTGTGATGCGCCGCTTC	390
Db	241	GGAACCGACGACGACCCGACATTTTCACATATGGCGGCCGCGTCCGTGTGATGCGCCGCTTC	300
Qy	391	ATGCGACCTTATTTTCGCTGCGTGGCGCGCCGAGGGGCTGCTGTCGCCGTATGTACGCTTAT	450
Db	301	ATGCGACCTTATTTTCGCTGCGTGGCGCGCGAGGGGCTGCTGTCGCCGTATGTACGCTTAT	360
Qy	451	GCGCTGATGTTTGGGAGGATCGCTGGATGATGATGATCTTCTGGCGGTTGCGTGCATCTG	510
Db	361	GCGCTGATCTTTGGGAGGATCGCTGGATGATGATGATCTTCTGGCGGTTGCGTGCATCTG	420
Qy	511	GCGTCGATCCAGCTGTTGCTGTTTCGCGCATCTGACTGCCGACCGCCCGGACACGACGCG	570
Db	421	GCGTCGATCCAGCTGTTGCTGTTTCGCGCATCTGACTGCCGACCGCCCGGACACGACGCG	480
Qy	571	TTCCCCGACCGCCACAAATGGCGCGTGTGTGGCGGATCAAGGACCCCGTGTGCTGTACCC	630
Db	481	TTCCCCGACCGCCACAAATGGCGCGTGTGTGGCGGATCAAGGACCCCGTGTGCTGTACCC	540
Qy	631	TGCTTTTCACTTTGGCGGTTTATCATCAGAAACACACTGCACCCGACCGTGTCTTGGTGG	690
Db	541	TGCTTTTCACTTTGGCGGTTTATCATCAGAAACACACTGCACCCGACCGTGTCTTGGTGG	600
Qy	691	CGCCTGCCGACACCCGACCAAGGGGGACACCGCATGA 729	
Db	601	CGCCTGCCGACACCCGACCAAGGGGGACACCGCATGA 639	
RESULT 8			
US-09-006-491-1			
Sequence 1, Application US/0906491			
Patent No. 5972690			
GENERAL INFORMATION:			
APPLICANT: MISAWA, No. 5972690ihiko			
APPLICANT: KONDO, Keiji			
APPLICANT: KAJIWARA, Susumu			
APPLICANT: YOKOYAMA, Akihito			
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF			
TITLE OF INVENTION: XANTHOPYLLS AND THE PROCESS FOR PRODUCING THE			
NUMBER OF SEQUENCES: 12			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Foley & Lardner			
STREET: 3000 K Street, N.W., Suite 500			
CITY: Washington			
STATE: D C.			
COUNTRY: USA			
ZIP: 20007-5109			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/006,491			

1	FLILING DATE:	
2	CLASSIFICATION:	
3	PRIOR APPLICATION DATA:	
4	APPLICATION NUMBER:	US 08/663,310
5	FILING DATE:	23-SEP-1996
6	APPLICATION NUMBER:	WO PCT/JP94/02220
7	FILING DATE:	26-DEC-1994
8	PRIOR APPLICATION DATA:	
9	APPLICATION NUMBER:	JP 6-235917
10	FILING DATE:	05-SEP-1994
11	PRIOR APPLICATION DATA:	
12	APPLICATION NUMBER:	JP 5-348737
13	FILING DATE:	27-DEC-1993
14	ATTORNEY/AGENT INFORMATION:	
15	NAME:	BENT, Stephen A.
16	REGISTRATION NUMBER:	29,768
17	REFERENCE/DOCKET NUMBER:	49441/109
18	TELECOMMUNICATION INFORMATION:	
19	TELEPHONE:	(202)672-5300
20	TELEFAX:	(202)672-5399
21	TELEX:	904136
22	INFORMATION FOR SEQ ID NO:	1:
23	SEQUENCE CHARACTERISTICS:	
24	LENGTH:	639 base pairs
25	TYPE:	nucleic acid
26	STRANDEDNESS:	double
27	TOPOLOGY:	linear
28	FEATURE:	
29	NAME/KEY:	CDS
30	LOCATION:	1..636
31	FEATURE:	
32	NAME/KEY:	mat_peptide
33	LOCATION:	1..636
34	US-09-006-491-1	
35	Query Match	83.9%; Score 611.8; DB 2; Length 639;
36	Best Local Similarity	97.3%; Pred. No. 1.6e-138;
37	Matches 622; Conservative	0; Mismatches 17; Indels 0; Gaps 0
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118		
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191</		

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5339
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 729 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..726
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..726
US-08-663-310-8

Query Match 59.3%; Score 432.4; DB 1; Length 729;
Best Local Similarity 77.6%; Pred. No. 3,6e-95;
Matches 523; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

56 CGGGGGGATCATCGCGCGGCTGGCGCGCTGATGATCGCGCTGCTTGGAGC 115
56 CGCGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 115
116 CGCGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175
116 CGCGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175
176 GTCTGTTATCATCGCGATGACGCGATGATGATGATGATGATGATGATGATG 235
176 GCTGTTATCATCGCGATGACGCGATGATGATGATGATGATGATGATGATG 235
236 CCAATCGCGGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCG 295
236 CCAATCGCGGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCG 295
236 TGAATGTCAGACATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 355
236 TGAATGTCAGACATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 355
236 TGAATGTCAGACATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 355
356 ACCATGCGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 415
356 GTACGAGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 415
416 AGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
416 AGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
476 TGTATGTCATCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
476 TGTATGTCATCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
536 GCATCTGCTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 595

536 GAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
596 CGTGGCGATGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 655
596 CGTGGCGATGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 655
656 ACGAACACACCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 715
656 ACGAACACACCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 715
716 GAGGCGCGCGATGA 729
716 GAGGCGCGCGATGA 729

RESULT 11
US-09-006-491-8
Sequence 8, Application US/09006491

GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5972690ihko
APPLICANT: KAJIWARA, Susumu
APPLICANT: KAJIWARA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5339
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 729 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..726

FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..726
US-09-006-491-8

Query Match 59.3%; Score 432.4; DB 2; Length 729;
Best Local Similarity 77.6%; Pred. No. 3.6e-95;
Matches 523; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 56 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 115
DB 56 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 115
QY 116 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 175
DB 116 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 175
QY 176 GTCGTTTCATCATCGCGCATGAGCGGATGCGATGCGGCTGCGCCCTGCGATGCGG 235
DB 176 GTCGTTTCATCATCGCGCATGAGCGGATGCGATGCGGCTGCGCCCTGCGATGCGG 235
QY 236 CCAATGCGGCGATGCGGCGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
DB 236 CCAATGCGGCGATGCGGCGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
QY 296 TGATGCGCAAGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGG 355
DB 296 TGATGCGCAAGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGG 355
QY 356 ACCATGCGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 415
DB 356 GTCACGCGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 415
QY 416 AGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
DB 416 AGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
QY 476 TGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535
DB 476 TGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535
QY 536 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
DB 536 GAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
QY 596 CGTGGCGGATGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
DB 596 CGACCGGCGATGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
QY 656 AGCAACACG 715
DB 656 AGCAACACG 715
QY 716 GGGACACCGCATGA 729
DB 716 GAGCGCGCGCATGA 729

RESULT 12
US-09-335-919-8
Sequence 8, Application US/09335919
Patent No. 6150130
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 6150130.hiko
APPLICANT: KONO, Keiji
APPLICANT: KAJIWARA, Suenmu
APPLICANT: YOROVAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,919
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 729 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..726
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..726
US-09-335-919-8

Query Match 59.3%; Score 432.4; DB 3; Length 729;
Best Local Similarity 77.6%; Pred. No. 3.6e-95;
Matches 523; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 56 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 115
DB 56 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 115
QY 116 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 175
DB 116 CGGCGCGCATCATCGCGCGCTGCGCCCTGCGATGCGATGCGGCTGCTTGTGAGC 175
QY 176 GTCGTTTCATCATCGCGCATGAGCGGATGCGATGCGGCTGCGCCCTGCGATGCGG 235
DB 176 GTCGTTTCATCATCGCGCATGAGCGGATGCGATGCGGCTGCGCCCTGCGATGCGG 235
QY 236 CCAATGCGGCGATGCGGCGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
DB 236 CCAATGCGGCGATGCGGCGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
QY 296 TGATGCGCAAGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGG 355
DB 296 TGATGCGCAAGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGGCGCATGCGG 355
QY 356 ACCATGCGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 415

```

Db      356 GTCAAGAGAGGCGCCGTGCGTGTACAGGAGCTTCTGTCCACTATTTTGGCGTGGCAG 415
QY      416 AGGGGCTGTGCTGCCCCGTCATCGTACGCGTCTATGCGGTATGTTGGGGATCGCTGA 475
Db      416 AGGGAGCTGTGCTACCGGTGATGTACACCACTATGCGTATCTTGAGCGATCCCTGA 475
QY      476 TGTACGTGTCTTCTGCGCCGTGCGTGTACCTGTGCGTGTATCCAGCTGTGTGTTCG 535
Db      476 TGTATGTATCTTCTGCGCCGTGCGTGTACCTGTGCGTGTATCCAGCTGTGTGTTCG 535
QY      536 GCATCTGCTGCGCGACCGCCCGCGCGACGAGCGTTCCCGGACCGGCACATGCGCGGT 595
Db      536 GAACCTGTGCTGCGCCCGCGCGCGACGATGACGATTTTCCGACCGGACGAGGAGT 595
QY      596 CGTGGCGATGACGCGACCGCGTGTGCTGTGACCTGTTTCACTTTTGGCGGTATTC 655
Db      596 CGACCGGATCGCGCGACCGCGTGTGCTGTGACCTGTTTCACTTTTGGCGGTATTC 655
QY      656 AGGAACACCACTGCAACCGCGAGCGTGTGCGTGTGCGCGCGCGACCGCGCGACGAG 715
Db      656 AGGAACATCATCTGATCGCGATGCGCGTGTGCGTGTGCGCGCGCGCGACGAGAGCG 715
QY      716 GGGACACCGCATGA 729
Db      716 GAGCGCGCGCATGA 729

```

RESULT 13

```

US-08-663-310-12
; Sequence 12, Application US/08663310
; Patent No. 5811273
; GENERAL INFORMATION:
; APPLICANT: MISAWA, No. 5811273jhiko
; APPLICANT: KONDO, Keiji
; APPLICANT: KAJIWARA, Susumu
; APPLICANT: YOKOYAMA, Akihito
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/663,310
; FILING DATE: 23-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/02220
; FILING DATE: 26-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-235917
; FILING DATE: 05-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-348737
; FILING DATE: 27-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136

```

```

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1631 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-663-310-12

```

```

Query Match      59.3%; Score 432.4; DB 1; Length 1631;
Best Local Similarity 77.6%; Pred. NO. 4.3e-95;
Matches 523; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```

```

QY      56 CGGGGCGATCATGCGCGCGTGGCTGGCTGCGATGTCATGCGCTGTGTTCTGACG 115
Db      154 CGCGCGCATCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 213
QY      116 CGGGCGCGCATCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 175
Db      214 CGGGCGCGCATCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 273
QY      176 GTCTGTTCATGATCGCGCATGACGCGATGACATGAGTGGTGGTGGTGGTGGTGGTGGTGG 235
Db      274 GGTCTTTCATGATCGCGCATGACGCGATGACATGAGTGGTGGTGGTGGTGGTGGTGGTGG 333
QY      236 CCAATGCGCGCATGAGCGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 295
Db      334 CCAATGCGCGCATGAGCGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 393
QY      296 TGATGTCACGACATGAGCGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 355
Db      394 TGATGTCACGACATGAGCGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 453
QY      356 ACCATGCGCGCGCGTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 415
Db      454 GTCAAGAGAGGCGCCGTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 513
QY      416 AGGGGCTGTGCTGCGCGTCAATGTCAGCGTGTATGCGCTGTATGTTGGGGATCGCTGA 475
Db      514 AGGGAGCTGTGCTGCGCGTCAATGTCAGCGTGTATGCGCTGTATGTTGGGGATCGCTGA 573
QY      476 TGTACGTGTCTTCTGCGCCGTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 535
Db      574 TGTATGTATCTTCTGCGCCGTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 633
QY      536 GCATCTGCTGCGCGACCGCCCGCGCGACGAGCGTTCGCGACCGCACCATATGCGCGGT 595
Db      634 GAACCTGTGCTGCGCGACCGCCCGCGCGACGAGCGTTCGCGACCGCACCATATGCGCGGT 693
QY      596 CGTGGCGATGACGCGACCGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 655
Db      694 CGACCGGATCGCGCGACCGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 753
QY      656 AGGAACACCACTGCAACCGCGAGCGTGTGCGTGTGCGCGCGCGACCGCGCGACGAG 715
Db      754 AGGAACATCATCTGATCGCGATGCGCGTGTGCGTGTGCGCGCGCGACGAGAGCG 813
QY      716 GGGACACCGCATGA 729
Db      814 GAGCGCGCGCATGA 827

```

RESULT 14

```

US-09-006-491-12
; Sequence 12, Application US/09006491
; Patent No. 5972690
; GENERAL INFORMATION:
; APPLICANT: MISAWA, No. 5972690jhiko
; APPLICANT: KONDO, Keiji
; APPLICANT: KAJIWARA, Susumu
; APPLICANT: YOKOYAMA, Akihito
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE

```

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1631 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-006-491-12

Query Match 59.3%; Score 432.4; DB 2; Length 1631;
Best Local Similarity 77.6%; Pred. No. 4.3e-95;
Matches 523; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

56 CGGGCGGCATCGCGCGGCTGAGCGGCTGATGCGATGCGGCTGAGCG 115
154 CGGGCGGCATCGCGCGGCTGAGCGGCTGATGCGATGCGGCTGAGCG 213
116 CGGGCGGCATCGCGCGGCTGAGCGGCTGATGCGATGCGGCTGAGCG 175
214 CGGGCGGCATCGCGCGGCTGAGCGGCTGATGCGATGCGGCTGAGCG 273
176 GTCGTTCATCATCGCGGCGATGAGCGGATGATGAGCGGCGCGCGCG 235
274 GCGTGTTCATCATCGCGGCGATGAGCGGATGATGAGCGGCGCGCGCG 333
236 CCAATCGCGGATGAGCGGCGATGAGCGGATGATGAGCGGCGCGCGCG 295
334 CCAATCGCGGATGAGCGGCGATGAGCGGATGATGAGCGGCGCGCGCG 393
296 TGAATCGCAAGCATGAGCGGCGATGAGCGGATGATGAGCGGCGCGCG 355
394 TGAATCGCAAGCATGAGCGGCGATGAGCGGATGATGAGCGGCGCGCG 453
356 ACATGCGCGCGCGCGCGCGGCTGAGCGGCGGCTTCATCGGCACTATT 415
454 GTTACCGAGAGCGCGCGCGGCTGAGCGGCGGCTTCATCGGCACTATT 513
416 AGGGCGTGTGCTGCGCGGCTGATGAGCGGCTGATGAGGCGATCGCTGA 475

514 AGGACTGTGCTACCGGTGATGATGATGATGATGATGATGATGATGAT 573
476 TGAATCGCAAGCATGAGCGGCGATGAGCGGATGATGAGCGGCGCGCG 535
574 TGAATCGCAAGCATGAGCGGCGATGAGCGGATGATGAGCGGCGCGCG 633
536 GCATGCTGCGCGCGCGCGCGCGGCGCGCGGCGCGCGGCGCGCGCG 595
634 GAATGCTGCGCGCGCGCGCGCGGCGCGCGGCGCGCGGCGCGCGCG 693
596 CGTGGCGATGAGCGGCGCGCGGCGCGGCGCGGCGCGGCGCGGCG 655
694 CGACCGGATGAGCGGCGCGCGGCGCGGCGCGGCGCGGCGCGGCG 753
656 ACAGACACGATGAGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCG 715
754 ACAGACACGATGAGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCG 813
716 GGGACACCGCATGA 729
814 GAGCGCGCGCATGA 827

RESULT 15
US-09-335-919-12
Sequence 12, Application US/09335919
Patent No. 6150130
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 61501301h1ko
APPLICANT: KONDO, Keiji
APPLICANT: KAJIWARA, Susumu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,919
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 12:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1631 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-335-919-12

```

Query Match	59.3%	Score	432.4	DB	3	Length	1631
Best Local Similarity	77.6%	Pred	No. 4.3e-95				
Matches	523	Conservative	0	Mismatches	151	Indels	0
						Gaps	0

[illegible]

Search completed: June 22, 2005, 01:09:51
Job time : 198.419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: June 21, 2005, 22:57:55 ; Search time 524.43 Seconds
(without alignments)
8628.994 Million cell updates/sec

Title: US-10-695-980-31

Perfect score: 729

Sequence: 1 atagcgcacacatcctcgtcc.....ccaagggggacacccgatga 729

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 6054689 seqs, 310372919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	729	100.0	729	15	US-10-166-225A-180
2	729	100.0	729	18	US-10-695-980-31
3	729	100.0	1261	10	US-09-920-923-30
4	729	100.0	1261	18	US-10-695-980-30
5	697	95.6	729	22	US-10-997-844-31
6	533.8	73.2	756	22	US-10-997-844-7
7	383.8	52.6	726	9	US-09-547-267-11

8	383.8	52.6	726	10	US-09-920-923-28	Sequence 28, Appl
9	383.8	52.6	726	18	US-10-695-980-28	Sequence 28, Appl
10	217	29.8	777	14	US-10-166-037-1	Sequence 1, Appl
11	195.4	26.8	735	18	US-10-466-656-38	Sequence 38, Appl
12	195.4	26.8	735	21	US-10-848-307-38	Sequence 38, Appl
13	69.2	9.5	1761	9	US-09-791-687A-1	Sequence 1, Appl
14	69.2	9.5	1761	9	US-09-791-687A-2	Sequence 2, Appl
15	69.2	9.5	1761	9	US-09-791-687A-3	Sequence 3, Appl
16	66.2	9.1	1458	10	US-09-920-923-31	Sequence 31, Appl
17	58.8	8.1	825	17	US-10-369-493-40591	Sequence 40591, A
18	52.4	7.2	1038	17	US-10-310-154-95	Sequence 95, Appl
19	52.4	7.2	1038	21	US-10-732-923-210	Sequence 210, Appl
20	52.4	7.2	1091	21	US-10-947-979-47	Sequence 47, Appl
21	51.2	7.0	522	15	US-10-156-761-3323	Sequence 3323, Appl
22	51.2	7.0	9025608	15	US-10-156-761-1	Sequence 1, Appl
23	50.4	6.9	1189	17	US-10-260-238-3770	Sequence 3770, Appl
24	49.8	6.8	741	15	US-10-156-761-4808	Sequence 4808, Appl
25	49.8	6.8	1431	17	US-10-282-122A-7249	Sequence 7249, Appl
26	49.8	6.8	9025608	15	US-10-156-761-1	Sequence 1, Appl
27	49.8	6.7	3131	17	US-10-282-122A-7241	Sequence 7241, Appl
28	49.8	6.7	13029	9	US-09-815-242-4052	Sequence 4052, Appl
29	49.8	6.7	13029	17	US-10-282-122A-7240	Sequence 7240, Appl
30	48.6	6.7	1243	18	US-10-425-114-24278	Sequence 24278, A
31	48.6	6.7	2031	17	US-10-369-493-42813	Sequence 42813, A
32	48.6	6.6	1618	18	US-10-425-114-35569	Sequence 35569, A
33	48.6	6.6	1732	20	US-10-425-115-109983	Sequence 109983, A
34	48.6	6.6	3183	17	US-10-282-122A-14200	Sequence 14200, A
35	47.4	6.5	3178	19	US-10-437-963-47749	Sequence 47749, A
36	47.4	6.4	873	20	US-10-425-115-183240	Sequence 183240, A
37	46.6	6.4	2421	20	US-10-425-115-145240	Sequence 145240, A
38	46.2	6.3	669	19	US-10-437-963-44697	Sequence 44697, A
39	46.2	6.3	2002	19	US-10-437-963-68831	Sequence 68831, A
40	46.2	6.3	798	15	US-10-156-761-30337	Sequence 30337, Appl
41	46.2	6.3	2238	19	US-10-437-963-97363	Sequence 97363, A
42	46.2	6.3	10592	15	US-10-156-761-414	Sequence 414, Appl
43	45.8	6.3	100000	21	US-10-156-761-15103	Sequence 15103, A
44	45.8	6.3	1102	21	US-10-947-979-7	Sequence 7, Appl
45	45.8	6.3	1102	21	US-10-947-979-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-10-166-225A-180

; Sequence 180, Application US/10166225A

; Publication No. US20030148416A1

; GENERAL INFORMATION:

; APPLICANT: BERRY, Alan

; APPLICANT: BREITZEL, Warner

; APPLICANT: HUMBERT, Markus

; APPLICANT: LOPEZ-DILIBARTI, Rual

; APPLICANT: MAYER, Anne F.

; APPLICANT: YELISEEV, Alexei A.

; TITLE OF INVENTION: IMPROVED ISOPRENOL PRODUCTION

; FILE REFERENCE: C38435/121966

; CURRENT APPLICATION NUMBER: US/10/166,225A

; CURRENT FILING DATE: 2002-06-05

; NUMBER OF SEQ ID NOS: 197

; SOFTWARE: Patent version 3.1

; SEQ ID NO 180

; LENGTH: 729

; TYPE: DNA

; ORGANISM: Paracoccus carotinifaciens F-396

; NAME/KEY: CDS

; LOCATION: (1)..(726)

; OTHER INFORMATION: Beta-carotene Beta-4 oxygenase

US-10-166-225A-180

Query Match 100.0%; Score 729; DB 15; Length 729;
Best Local Similarity 100.0%; Pred. No. 6.3e-197;
Matches 729; Conservative 0; Mismatches 0; Gaps 0;

```
QY      1 ATAGGCGACATGCGCTGGCCCAAGGAGATGTGACCGGACAGTTGATCGTTCGGGC 60
      1 ATAGAGGCGACATGCGCTGGCCCAAGGAGATGTGACCGGACAGTTGATCGTTCGGGC 60
Db
QY      61 GGCATCATGCGCGCGTGGCTGGCCCTGATGATGTGATGCGCTGTGTTCTGGAACGGCGCG 120
      61 GGCATCATGCGCGCGTGGCTGGCCCTGATGATGTGATGCGCTGTGTTCTGGAACGGCGCG 120
Db
QY      121 GGGCATATCCATCTGGCGGCTGGCGAATTTCTGGGGGCTGACCGGCTGTGCGGTCTG 180
      121 GGGCATATCCATCTGGCGGCTGGCGAATTTCTGGGGGCTGACCGGCTGTGCGGTCTG 180
Db
QY      181 TTTCATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
      181 TTTCATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db
QY      241 GGGCGCATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
      241 GGGCGCATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db
QY      301 GTCAAGCAGATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
      301 GTCAAGCAGATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db
QY      361 GCGCGCGCGCGTGGCTGGTATGACCGCGCTTTCATGCGGACCTATTTTGGCTGGCGGAGGG 420
      361 GCGCGCGCGCGTGGCTGGTATGACCGCGCTTTCATGCGGACCTATTTTGGCTGGCGGAGGG 420
Db
QY      421 CTGCTGCTGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
      421 CTGCTGCTGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db
QY      481 GTGGTCTTCTGGCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 540
      481 GTGGTCTTCTGGCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 540
Db
QY      541 TGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
      541 TGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db
QY      601 CGGATCAGCGACCGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
      601 CGGATCAGCGACCGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
Db
QY      661 CACCACTGCAACCGCGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 720
      661 CACCACTGCAACCGCGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 720
Db
QY      721 ACCGCATGA 729
      721 ACCGCATGA 729
Db
RESULT 2
US-10-695-980-31
; Sequence 31, Application US/10695980
; Publication No. US20040058410A1
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; TITLE OF INVENTION: Fermentative Carotenoid Production
; FILE REFERENCE: 15464 US (38435/125944)
; CURRENT APPLICATION NUMBER: US/10/695,980
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/920,923B
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 31
; LENGTH: 729
```

```

; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: E-396
US-10-695-980-31
Query Match      100.0%; Score 729; DB 18; Length 729;
Best Local Similarity 100.0%; Pred. No. 6.3e-197;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATAGGCGACATGCGCGTGGCCCAAGGAGATGTGACCGGACAGTTGATCGTTCGGGC 60
      1 ATAGGCGACATGCGCGTGGCCCAAGGAGATGTGACCGGACAGTTGATCGTTCGGGC 60
Db
QY      61 GGCATCATGCGCGCGTGGCTGGCCCTGATGATGTGATGCGCTGTGTTCTGGAACGGCGCG 120
      61 GGCATCATGCGCGCGTGGCTGGCCCTGATGATGTGATGCGCTGTGTTCTGGAACGGCGCG 120
Db
QY      121 GGGCATATCCATCTGGCGGCTGGCGAATTTCTGGGGGCTGACCGGCTGTGCGGTCTG 180
      121 GGGCATATCCATCTGGCGGCTGGCGAATTTCTGGGGGCTGACCGGCTGTGCGGTCTG 180
Db
QY      181 TTTCATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
      181 TTTCATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db
QY      241 GGGCGCATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
      241 GGGCGCATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db
QY      301 GTCAAGCAGATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
      301 GTCAAGCAGATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db
QY      361 GCGCGCGCGCGTGGCTGGTATGACCGCGCTTTCATGCGGACCTATTTTGGCTGGCGGAGGG 420
      361 GCGCGCGCGCGTGGCTGGTATGACCGCGCTTTCATGCGGACCTATTTTGGCTGGCGGAGGG 420
Db
QY      421 CTGCTGCTGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
      421 CTGCTGCTGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db
QY      481 GTGGTCTTCTGGCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 540
      481 GTGGTCTTCTGGCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 540
Db
QY      541 TGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
      541 TGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db
QY      601 CGGATCAGCGACCGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
      601 CGGATCAGCGACCGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
Db
QY      661 CACCACTGCAACCGCGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 720
      661 CACCACTGCAACCGCGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 720
Db
QY      721 ACCGCATGA 729
      721 ACCGCATGA 729
Db
RESULT 3
US-09-920-923-30
; Sequence 30, Application US/09920923
; Publication No. US2003002273A1
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/09/920,923
```

;; CURRENT FILING DATE: 2001-08-02
;; PRIOR APPLICATION NUMBER: 08/980,832
;; PRIOR FILING DATE: 1997-12-01
;; NUMBER OF SEQ ID NOS: 66
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 30
;; LENGTH: 1261
;; TYPE: DNA
;; ORGANISM: Alcaligenes PC-1
US-09-920-923-30

Query Match 100.0%; Score 729; DB 10; Length 1261;
Best Local Similarity 100.0%; Pred. No. 6.7e-197;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGCGCACATGCTCTGCGCCCAAGGAGATCTGACCGCCACCGATTGATCGTCTGGGC 60
DB 40 ATGAGCGCACATGCTCTGCGCCCAAGGAGATCTGACCGCCACCGATTGATCGTCTGGGC 99
QY 61 GGCATATGCGCGCGCTGGGCTGGGCTGCAATGTCATGCGCTGATGTTCTGAGCGGGCG 120
DB 100 GGCATATGCGCGCGCTGGGCTGGGCTGCAATGTCATGCGCTGATGTTCTGAGCGGGCG 159
QY 121 GGCATATGCGCGCGCTGGGCTGGGCTGCAATGTTCTGAGGGCTGACCTGCTGCTGCTG 180
DB 160 GGCATATGCGCGCGCTGGGCTGGGCTGCAATGTTCTGAGGGCTGACCTGCTGCTGCTG 219
QY 181 TTTCATATGCGCGATGACCGGATGATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 220 TTTCATATGCGCGATGACCGGATGATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTG 279
QY 241 GCGGCGATGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 280 GCGGCGATGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 339
QY 301 GTCAAGCATGAGCCCATCATGCGCATGCGGAAACCGACGACCGACCAATTTCACAT 360
DB 340 GTCAAGCATGAGCCCATCATGCGCATGCGGAAACCGACGACCGACCAATTTCACAT 399
QY 361 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 400 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459
QY 421 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 460 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 481 GTGCTCTTGTGCGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG 540
DB 520 GTGCTCTTGTGCGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG 579
QY 541 TGGCTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 580 TGGCTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639
QY 601 CGGATAGGAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 640 CGGATAGGAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 699
QY 661 CACCACTTGACCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 700 CACCACTTGACCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
QY 721 ACCGATGA 729
DB 760 ACCGATGA 768
```

RESULT 4
US-10-695-980-30
; Sequence 30, Application US/10695980
; Publication No. US20040058410A1
; GENERAL INFORMATION:

;; APPLICANT: Pasamontes, Luis
;; APPLICANT: Teyankov, Yuri
;; TITLE OF INVENTION: Fermentative Carotenoid Production
;; FILE REFERENCE: 15464 US (C38435/125944)
;; CURRENT APPLICATION NUMBER: US/10/695,980
;; CURRENT FILING DATE: 2003-10-29
;; PRIOR APPLICATION NUMBER: US/09/920,923B
;; PRIOR FILING DATE: 2001-08-02
;; PRIOR APPLICATION NUMBER: 08/980,832
;; PRIOR FILING DATE: 1997-12-01
;; NUMBER OF SEQ ID NOS: 66
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 30
;; LENGTH: 1261
;; TYPE: DNA
;; ORGANISM: Alcaligenes PC-1
US-10-695-980-30

Query Match 100.0%; Score 729; DB 18; Length 1261;
Best Local Similarity 100.0%; Pred. No. 6.7e-197;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGCGCACATGCTCTGCGCCCAAGGAGATCTGACCGCCACCGATTGATCGTCTGGGC 60
DB 40 ATGAGCGCACATGCTCTGCGCCCAAGGAGATCTGACCGCCACCGATTGATCGTCTGGGC 99
QY 61 GGCATATGCGCGCGCTGGGCTGGGCTGCAATGTCATGCGCTGATGTTCTGAGCGGGCG 120
DB 100 GGCATATGCGCGCGCTGGGCTGGGCTGCAATGTCATGCGCTGATGTTCTGAGCGGGCG 159
QY 121 GGCATATGCGCGCGCTGGGCTGGGCTGCAATTTCTGAGGGCTGACCTGCTGCTGCTGCTG 180
DB 160 GGCATATGCGCGCGCTGGGCTGGGCTGCAATTTCTGAGGGCTGACCTGCTGCTGCTGCTG 219
QY 181 TTTCATATGCGCGATGACCGGATGATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 220 TTTCATATGCGCGATGACCGGATGATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 279
QY 241 GCGGCGATGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 280 GCGGCGATGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 339
QY 301 GTCAAGCATGAGCCCATCATGCGCATGCGGAAACCGACGACCGACCAATTTCACAT 360
DB 340 GTCAAGCATGAGCCCATCATGCGCATGCGGAAACCGACGACCGACCAATTTCACAT 399
QY 361 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 400 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459
QY 421 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 460 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 481 GTGCTCTTGTGCGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG 540
DB 520 GTGCTCTTGTGCGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG 579
QY 541 TGGCTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 580 TGGCTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639
QY 601 CGGATAGGAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 640 CGGATAGGAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 699
QY 661 CACCACTTGACCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 700 CACCACTTGACCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
QY 721 ACCGATGA 729
DB 760 ACCGATGA 768
```


Db 442 CTGTGTGTCGCGATCGTCACTGCTATGCGCTGATCTGGGCGACGCGTGAATAT 501
QY 481 GTGCTCTTCTGCGCGTTCGCGTGAATCTGCGCTGATCGAGCTGTTCGTTCGCGATC 540
Db 502 GTGTCTTCTGCGCGTTCGCGATCTGCGCGATCGAGCTGTTCGTTCGTTCGCGATC 561
QY 541 TGGCTGCGCGACCG 600
Db 562 TGGCTGCGCGATCG 621
QY 601 CGGATCAGCGACCG 660
Db 622 CGGATCAGCGACCG 681
QY 661 CACCATCTGACCG 720
Db 682 CATCATCTGATCG 741
QY 721 ACCGCGTGA 729
Db 742 ACCGCGTGA 750

RESULT 7

US-09-547-267-11
Sequence 11, Application US/09547267
Patent No. US20020147371A1

GENERAL INFORMATION:

APPLICANT: Hohmann, Hans-Peter
APPLICANT: Pasamontes, Luis
APPLICANT: Teasler, Michel
APPLICANT: van Loon, Adolphus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSER: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/547,267

FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: RAN 6002/170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-547-267-11

Query Match 52.6%; Score 383.8; DB 9; Length 726;
Best Local Similarity 74.8%; Pred. No. 6,1e-99;
Matches 481; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 76 TGGTGGCCCTGATATGATGATGCGCTGTGATTTCTGACGCGCGCGCGCATCCATCTG 135
Db 76 TGGTGGCTTCTGACGCTTTCACTGTGTCTGTGATGATGATGATGATGATGATGAT 135
QY 136 GCGGTGCGAATTTCTGCGCGCTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 195
Db 136 GCTGTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 195
QY 196 GAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
Db 196 GAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
QY 256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 315
Db 256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 315
QY 316 CATCATCGCATGCG 375
Db 316 CACCATCGCATGCG 375
QY 376 TGGTACGCGCTTCTGATCGCGCATTTTCTGCGCGCGCGCGCGCGCGCGCGCGCG 435
Db 376 TGGTACGCGCTTCTGATCGCGCATTTTCTGCGCGCGCGCGCGCGCGCGCGCGCG 435
QY 436 ATGATGCGGCTGATGCG 495
Db 436 ATGATGCGGCTGATGCG 495
QY 496 TTGCGCTGATCTGCG 555
Db 496 GTTCCGCGCTGTCTGTGCTTCTGATCGCGCATTTTCTGCGCGCGCGCGCGCGCGCG 555
QY 556 CCGCGCGACGAGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 615
Db 556 CCGCGCGACGAGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 615
QY 616 GTGTGCTGCTGACCGCTTTCACCTTTGCGCGGTATCATCAAGACGCGCGCGCGCG 675
Db 616 GTGTGCTGCTGACCGCTTTCACCTTTGCGCGGTATCATCAAGACGCGCGCGCGCG 675
QY 676 ACCTGCTGCTGTGTGCG 718
Db 676 CAGCTTCGCTGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 718

RESULT 8

US-09-920-923-28
Sequence 28, Application US/09920923
Publication No. US20030022273A1

GENERAL INFORMATION:

APPLICANT: Pasamontes, Luis
APPLICANT: Teysankov, Yuri
TITLE OF INVENTION: Improved Fermentative Carotenoid Production
FILE REFERENCE: Improved Fermentative Carotenoid
CURRENT APPLICATION NUMBER: US/09/920,923
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 08/980,832
PRIOR FILING DATE: 1997-12-01
NUMBER OF SEQ ID NOS: 66
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 28
LENGTH: 726
TYPE: DNA
ORGANISM: Alcaligenes PC-1
US-09-920-923-28

Query Match 52.6%; Score 383.8; DB 10; Length 726;
Best Local Similarity 74.8%; Pred. No. 6,1e-99;
Matches 481; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

Db 76 TGGCTGCTTCTGACGCTTTACCTGTGGCTGTGACGCTGCTACCCGCTGCTG 135
136 GGGGTGCGAAATTTCTGAGGAGCTGACCTGTGCTGTGCTGTTCATTCATCGCAT 195
136 GCTGTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGAC 195
Qy 196 GACCGATGATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 255
196 GACCGATGATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 255
Db 196 GACCGATGATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 255
Qy 256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 315
256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 315
Db 256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 315
Qy 316 CATCATGCGCATGCGGAAACCGACGACGACGACGACGACGACGACGACGACGACGACG 375
316 CACCAACCGTACGCTGTGATCCGACGACGACGACGACGACGACGACGACGACGACG 375
Db 316 CACCAACCGTACGCTGTGATCCGACGACGACGACGACGACGACGACGACGACGACG 375
Qy 376 TGGTACGCGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 435
376 TGGTACGCGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 435
Db 376 TGGTACGCGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 435
Qy 436 ATCGTACGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 495
436 ATCGTACGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 495
Db 436 ATCGTACGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 495
Qy 496 TTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 555
496 TTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 555
Db 496 TTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 555
Qy 556 CCGGCGCACGACGCTTCCCGGACCGGACGACGACGACGACGACGACGACGACGACGACG 615
556 CCGGCGCACGACGCTTCCCGGACCGGACGACGACGACGACGACGACGACGACGACGACG 615
Db 556 CCGGCGCACGACGCTTCCCGGACCGGACGACGACGACGACGACGACGACGACGACGACG 615
Qy 616 GTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 675
616 GTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 675
Db 616 GTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 675
Qy 676 ACGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 718
676 ACGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 718
Db 676 ACGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 718

RESULT 9
US-10-695-980-28
; Sequence 28, Application US/10695980
; Publication No. US20040058410A1
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsyanikov, Yuri
; TITLE OF INVENTION: Fermentative Carotenoid Production
; FILE REFERENCE: 15464 US (C38435/125944)
; CURRENT APPLICATION NUMBER: US/10/695,980
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/920,923B
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; TYPE: DNA
; LENGTH: 726
; ORGANISM: Alicaligenes PC-1
US-10-695-980-28

Query Match 52.6%; Score 383.8; DB 18; Length 726;
Best Local Similarity 74.8%; Pred. No. 6.1e-99;
Matches 481; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
Qy 76 TGGCTGCGCTGTGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 135
76 TGGCTGCGCTGTGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 135
Db 76 TGGCTGCGCTGTGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 135

Qy 136 GCGGTGCGAAATTTCTGAGGAGCTGACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 195
136 GCTGTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGAC 195
Db 136 GCTGTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGAC 195
Qy 196 GACCGATGATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 255
196 GACCGATGATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 255
Db 196 GACCGATGATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 255
Qy 256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 315
256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 315
Db 256 CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 315
Qy 316 CATCATGCGCATGCGGAAACCGACGACGACGACGACGACGACGACGACGACGACGACG 375
316 CACCAACCGTACGCTGTGATCCGACGACGACGACGACGACGACGACGACGACGACG 375
Db 316 CACCAACCGTACGCTGTGATCCGACGACGACGACGACGACGACGACGACGACGACG 375
Qy 376 TGGTACGCGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 435
376 TGGTACGCGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 435
Db 376 TGGTACGCGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 435
Qy 436 ATCGTACGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 495
436 ATCGTACGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 495
Db 436 ATCGTACGCGCTTCAATCGGACCTATTTGCGCTGTGCGGAGGAGGCTGTGCTGCGCGCT 495
Qy 496 TTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 555
496 TTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 555
Db 496 TTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 555
Qy 556 CCGGCGCACGACGCTTCCCGGACCGGACGACGACGACGACGACGACGACGACGACGACG 615
556 CCGGCGCACGACGCTTCCCGGACCGGACGACGACGACGACGACGACGACGACGACGACG 615
Db 556 CCGGCGCACGACGCTTCCCGGACCGGACGACGACGACGACGACGACGACGACGACGACG 615
Qy 616 GTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 675
616 GTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 675
Db 616 GTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 675
Qy 676 ACGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 718
676 ACGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 718
Db 676 ACGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 718

RESULT 10
US-10-166-037-1
; Sequence 1, Application US/10166037
; Publication No. US2003008737A1
; GENERAL INFORMATION:
; APPLICANT: GIRAUD, Eric
; APPLICANT: HANNIBAL, Laure
; TITLE OF INVENTION: Isolated Carotenoid biosynthesis gene cluster involved
; FILE REFERENCE: 1721-55
; CURRENT APPLICATION NUMBER: US/10/166,037
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/297,272
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: crtlw
US-10-166-037-1

Query Match 29.8%; Score 217; DB 14; Length 777;
Best Local Similarity 61.3%; Pred. No. 1.3e-51;
Matches 425; Conservative 0; Mismatches 250; Indels 18; Gaps 4;
Qy 21 CAGGCAATGATGACCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 80
21 CAGGCAATGATGACCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 80
Db 51 CGATGCGAGGACGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 110

Matches 407; Conservative 0; Mismatches 256; Indels 18; Gaps 4

676 ACGTGCCCTGGTGGCCCTG 696
697 TGGCGGCCCTGGTGGCGTCTG 717

-307-38
 e 38, Application US/10848307
 tion No. US2005000347A1
 INFORMATION:
 INVENTOR: desouza, Mervyn L.
 INVENTOR: Schroeder, William A.
 INVENTOR: Kollmann, Sherry R.
 INVENTOR: May, Colleen A.
 OF INVENTION: Carotenoid Biosynthesis
 REFERENCE: CGL00/0243US02
 T APPLICATION NUMBER: US/10/848,307
 T FILING DATE: 2004-05-18
 APPLICATION NUMBER: US 10/466,656
 FILING DATE: 2003-07-18
 APPLICATION NUMBER: US 60/288,984
 FILING DATE: 2001-05-04
 APPLICATION NUMBER: US 60/264,329
 FILING DATE: 2001-01-26

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38

```

LENGTH: 735
 TYPE: DNA
 ORGANISM: Brevundimonas aurantiaca
 US-10-848-307-38

Query Match 26.8%; Score 195.4; DB 21; Length 735;
 Best Local Similarity 59.8%; Pred. No. 1.8e-45;
 Matches 407; Conservative 0; Mismatches 256; Indels 18; Gaps 4;

```

QY 31 CTGACCGCACAGATTGATCGTCGCGCGGCATCATCGCGCGGTGGCCCTGCAT 90
   |||||
DB 40 CAGACCTGATGCTGACCTCGCGGAGATGATGCGGAGTGGCGGTTCGAT 99
   |||||
QY 91 GGCATGCGCTGCTGCTTCTGACGCGCGCGCATCCCATCTGCGCGTCCGATTTTC 150
   |||||
DB 100 GCTCATGCGCTTATTTTCAACGATGCGGCGCGTTCGACCTGTGATATGCCCGGCGATC 159
   |||||
QY 151 CTGGAG---CTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
   |||||
DB 160 GTGGCGGCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 219
   |||||
QY 208 GGGTGGTGTGCGCGCGCGCGCGCGCGCATATGCGCGCATGCGCGCATGCTGCTGCTG 267
   |||||
DB 220 GACTCCCTGCGCGCGGAGCGCGCGCTGAAACGCGCATGCGCGCGCTGCGCGCTGCG 279
   |||||
QY 268 CTGTATGCGCGATTTTCTGCGCGCAAGATGATGCTCAAGCATGCGCGCATGCGCGCAT 327
   |||||
DB 280 CTCTATGCGGCTTCCGCTTGCATGCGCTGAAAGACGCGCGCATGCGCGCATGCGCGCAT 339
   |||||
QY 338 GCGGAAACGACGACGACCGCATGTTGAC-----CATGCGCGCGCGGTCCGCTG 378
   |||||
DB 340 CCGGACAGCGCGCGACCGCGGATTTTCAAGCGCGCGCGCGCGCTTCTTCCCTG 399
   |||||
QY 379 TACGCGCGCTCATGCGGACCTATTCGCGCGCGCGCGCGCGCGCTGCTGCGCGCATC 438
   |||||
DB 400 TTCCTGAATCTTCTTGCACCTATTTTGGCTGCGCGGAGATGCGCGCTGCGCGCGCTG 459
   |||||
QY 439 GTGACGCTCATGCGCTGATGTTGGGAGATCGCTGATGTA---GTTGCTTCTGCGCG 495
   |||||
DB 460 GTCCTGATGCGCTTCTTGGCGCGCTGCGCGCGCGCGCGCGCGCATCTCTGATCTTGGGCG 519
   |||||
QY 496 TTGCGCTGATCTTGGCGCTGATTCAGCTGTTGCTGCTGCGCATGCTGCTGCGCGCG 555
   |||||
DB 520 GCGCGCGCGCTTCTTCAAGCGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCG 579
   |||||
QY 556 CCGGCGACGACGCGCTTCCGCGACCGCGCATGCGCGGTGCTGCGCATGCGCGCGCG 615
   |||||
DB 580 CACACCGACGACCGCGCTTCCGCGACCGCGCATGCGCGCGCGCGCGCGCGCGCGCG 639
   |||||
QY 616 GTGCGCTGCTGACCTGCTTTCATCTTGGCGGTTATCATCAGAACACCACTGCGACCG 675
   |||||
DB 640 CTTTCCCTGCTCACTGCTTTCACCTTCCGCGCG---CCACACGAGAACACCATGAGCGCC 696
   |||||
QY 676 ACGTGCCTTGTGCGCGCGCTG 696
   |||||
DB 697 TGGCGCGCGCTGCTGCGCTG 717
   |||||

```

RESULT 13
 US-09-791-687A-1

Sequence 1, Application US/09791687A
 Patent No. US20020053096A1
 GENERAL INFORMATION:
 APPLICANT: Hirschberg, Joseph
 APPLICANT: Lotan, Tamar
 TITLE OF INVENTION: POLYNUCLEOTIDE MOLECULE FROM HAEMATOCOCCUS PLUVIALIS ENCODING A
 TITLE OF INVENTION: POLYPEPTIDE HAVING A BETA-C-4-OXYGENASE ACTIVITY FOR BIOTECHNOLO
 TITLE OF INVENTION: OF (3S,3'S) ASTAXANTHIN AND ITS SPECIFIC EXPRESSION IN CHROMOBLA
 TITLE OF INVENTION: PLANTS
 FILE REFERENCE: 01/21600
 CURRENT APPLICATION NUMBER: US/09/791,687A
 CURRENT FILING DATE: 2001-02-26
 PRIOR APPLICATION NUMBER: US 09/259,294

PRIOR FILING DATE: 1999-03-01
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: Patent in version 3.0
 SEQ ID NO 1
 LENGTH: 1761
 TYPE: DNA
 ORGANISM: Haematococcus pluvialis
 US-09-791-687A-1

Query Match 9.5%; Score 69.2; DB 9; Length 1761;
 Best Local Similarity 50.7%; Pred. No. 1.3e-09;
 Matches 223; Conservative 0; Mismatches 208; Indels 9; Gaps 2;

```

QY 139 GTCCGAATTTCTGCGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 198
   |||||
DB 526 GTCCGAATTTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585
   |||||
QY 199 GCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 258
   |||||
DB 586 GCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
   |||||
QY 259 GTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
   |||||
DB 646 TGCATCTCTTGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
   |||||
QY 319 CATGCGCA---TGCGGAAACGACGACGACCGCATGTTGACCATGCGCGCGCG----- 369
   |||||
DB 706 CACAAACGACGCTGCGGAGAGTGGGCAAGGACCGCTGCTTCCACAGGAGAAACCTTGGCAT 765
   |||||
QY 370 GTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
   |||||
DB 766 GTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825
   |||||
QY 430 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
   |||||
DB 826 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
   |||||
QY 490 TGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
   |||||
DB 886 ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 945
   |||||
QY 550 CACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 569
   |||||
DB 946 CACAGCCTGAGCGCTGCGCG 965
   |||||

```

RESULT 14
 US-09-791-687A-2

Sequence 2, Application US/09791687A
 Patent No. US20020053096A1
 GENERAL INFORMATION:
 APPLICANT: Hirschberg, Joseph
 APPLICANT: Lotan, Tamar
 TITLE OF INVENTION: POLYNUCLEOTIDE MOLECULE FROM HAEMATOCOCCUS PLUVIALIS ENCODING A
 TITLE OF INVENTION: POLYPEPTIDE HAVING A BETA-C-4-OXYGENASE ACTIVITY FOR BIOTECHNOLO
 TITLE OF INVENTION: OF (3S,3'S) ASTAXANTHIN AND ITS SPECIFIC EXPRESSION IN CHROMOBLA
 TITLE OF INVENTION: PLANTS
 FILE REFERENCE: 01/21600
 CURRENT APPLICATION NUMBER: US/09/791,687A
 CURRENT FILING DATE: 2001-02-26
 PRIOR APPLICATION NUMBER: US 09/259,294
 PRIOR FILING DATE: 1999-03-01
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: Patent in version 3.0
 SEQ ID NO 2
 LENGTH: 1761
 TYPE: RNA
 ORGANISM: Haematococcus pluvialis
 US-09-791-687A-2

Query Match 9.5%; Score 69.2; DB 9; Length 1761;
 Best Local Similarity 37.7%; Pred. No. 1.3e-09;
 Matches 166; Conservative 57; Mismatches 208; Indels 9; Gaps 2;

QY 139 ATCCGGATTTTCTCGGGGCTACCGGTGTCGGGTCTGTTTCATCATGCGCATGAC 198
 Db 526 GTCCGAGAAUUCUUUUGUCCUGGAGUUCUUGUACACAGGCCUUUUUUAUCCACGCAUGAU 585
 QY 199 GCGATGATCGAGTCGATCGTCGCGGGGCGCCCGCGCCGCAATGCGGCGATGAGCCAGCTT 258
 Db 586 GCUAUGCAUUGGCAACUCCAGCCCAUAGAAGCAGCUUAAUAGACUUUUGGGCAGAGUA 645
 QY 259 GTCCGTGAGCTGTATGCGCGATTTTCCGTGGGGCAAGATGATCTTCACACATGAGCCAT 318
 Db 646 UGCACUUCUUUGUAGCCUUGUUUAGUUAACAACAGUCUGCACCCGAAACAUUGGGAGCAC 705
 QY 319 CATGCGCA---TGCCGGAACCGACGACGACCCAGATTCGACATGCGCGGCC----- 369
 Db 706 CACAACCAACAUUGGAGAGUGGGCAAGGACCUUGACUUCGACAGGGGAAACCUUGCAUU 765
 QY 370 GTCCGCTGTATCGCCCGCTTCATCGGCACTTATTTCCGCTGGCGCGAGGGGCTGTGCTG 429
 Db 766 GUGCCUUGUUUGCCACGCUUCAUUGCCAGCAACUUGAGUUGGCAUUUUGCGCGCC 825
 QY 430 CCCGTCATCGTACGAGCTATATGCGCTGATGTTGGGGGAATCGCTGATGTACGTGATTTTC 489
 Db 826 GCAUGGUGAAGCGUGGUAUGCAGCUCUGCGUGGUGGCGCAAUUGGCAACUUCUGUGUUC 885
 QY 490 TGGCCGTTGCCGTGATCTTGAGCGCTGATCGAGCTGTTGTTGTTGGGATCTGAGTCGCG 549
 Db 886 AUGGCGGCGGCGCCCAUCCUUGCCGCCUUCGCUUUGUUAUUGGACCAUACUAGCCC 945
 QY 550 CACCGCCCCGGCCACAGAGC 569
 Db 946 CACAAGCCUAGACUUGGCGC 965

```

RESULT 15
US-09-791-687A-3
: Sequence 3, Application US/09791687A
: Patent No. US20020053096A1
: GENERAL INFORMATION:
: APPLICANT: Hirschberg, Joseph
: TITLE OF INVENTION: LOEAM, Tamar
: TITLE OF INVENTION: POLYNUCLEOTIDE MOLECULE FROM HAEMATOCOCCUS PLUVIALIS ENCODING A
: TITLE OF INVENTION: POLYPEPTIDE HAVING A BETA-C-4-OXYGENASE ACTIVITY FOR BIOCHEMOL
: TITLE OF INVENTION: OF (3S,3'S) ASTAXANTHIN AND ITS SPECIFIC EXPRESSION IN CHROMOPLA
: TITLE OF INVENTION: PLANTS
: FILE REFERENCE: 01/21600
: CURRENT APPLICATION NUMBER: US/09/791,687A
: CURRENT FILING DATE: 2001-02-26
: PRIOR APPLICATION NUMBER: US 09/259,294
: PRIOR FILING DATE: 1999-03-01
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3
: LENGTH: 1761
: TYPE: DNA
: ORGANISM: Haematococcus pluvialis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (166)..(1152)
: US-09-791-687A-3

```

Query Match	9.5%	Score 69.2;	DB 9;	Length 1761;
Best Local Similarity	50.7%	Pred. No. 1.3e-09;		
Matches 223; Conservative	0;	Mismatches 208;	Indels 9;	Gaps 2;

Accession	Sequence	Position
Qy	139 GTCGGCAATTTCTCTGGGCGCTGACCTGCGTCCGATCGATCTGTTTCATCATCGGCGCATGAC	138
Db	526 GTCGTAAGATTCTTTTCTCTGGAGTTCCGTACAACAGGCGCTTTTATTCACACAGCAGATGAT	585
Qy	139 GCGATGCATGGGTCCGTCGTGCGCGGGGCGCCCGCGCGCCCAATGCGGCGCATGGGCGACCTT	258
Db	586 GGTATGACATGGGACCATCGCATGAGAAAACAGGCAAGCTTATATACCTTCTTTGGGCAAGTA	645

QY	259	GTCCGTGAGCGTATGACCGGATTTTCTGCGGCAAGATGATGTCGAAGCAATGGCCCAT	318
Db	646	TGCATCTCTCTGTACGCTGTGTTTAATTAAACAATGCTGACCCGAAAGCATTTGGGAGCAC	705
QY	319	CATCGCCA---TGCCGGAAACGACGACGACCCAGATTTTCGACCATGGCGGCCG-----	369
Db	706	CACAAACCACTGCGGGAGGTGGGCAAGAACCTTGACTTCCACAGGGGAAACCTTGGCATT	765
QY	370	GTCCGCTGGTACGCCCGCTTCATCGGACCTTATTTGCGCTGCGCAGGGGCTGTGCTG	429
Db	766	GTGCCCTGGTTTGCAGCTTATGATTCACAGTCAATGTCATGTGCGAGATTGGCGCCCTC	825
QY	430	CCCGTCATCTGACGCTCTATGCGCTGATGTTGGGGGATCGCTGAGATGTAACGTGCTTC	489
Db	826	GCATGATGACCGCTGTGTCATGACGCTGTGCTGGGTGCGCCATGCGAACTGTCTGGTCTC	885
QY	490	TGGCGGTTGCGGTGCATCTCGGCGTTCGATTCACAGCTGTTCGTGTTGGGCATCTGGCTGCCG	549
Db	886	ATGGGGGCGCGCCCATCTGTGCGGCTTTCGCGCTGTCTTACTTTGGCACTACATGCCC	945
QY	550	CACCGCCCGCGGACGACGCGC	569
Db	946	CACAAAGCTTGAGCTTGCGCC	965

Search completed: June 22, 2005, 04:23:34
Job time : 534.43 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 19:42:50 ; Search time 2873.3 Seconds

(without alignments)
9657.484 Million cell updates/sec

Title: US-10-695-980-31

Perfect score: 729

Sequence: 1 atgagcgacacgacccctgc.....ccaagggggacaccgatga 729

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	8.9	666	4	BG846752 1024014H0
2	55.6	7.6	925	4	CNS00091P
3	52	7.1	1201	9	CNS01690
4	50.6	6.9	1101	9	CNS0178Y
5	50	6.9	925	9	CNS00091P
6	49	6.7	849	8	CC407345
7	48.8	6.7	427	5	BQ245798
8	48.8	6.7	663	6	CA108291
9	48.4	6.6	1100	9	CNS01690
10	48.2	6.6	633	6	CA161275
11	48.2	6.6	677	6	CA161511
12	48	6.6	613	7	COS33928
13	48	6.6	785	6	CG236345
14	47.6	6.5	634	6	CA096815
15	47.2	6.5	933	4	BG343324
16	47.2	6.5	968	8	AG030603
17	47.2	6.5	1169	8	BZ561896
18	47.2	6.5	1206	9	CL971023
19	47	6.4	693	7	COS20246
20	47	6.4	906	9	CG354638
21	47	6.4	907	9	CG354627
22	46.6	6.4	946	9	CG290659
23	46.6	6.4	351	4	B1479111
24	46.6	6.4	658	6	CA247911

C 25	46.6	6.4	664	6	CA119890	CCCLR106
C 26	46.6	6.4	1101	9	CNS01258	AL101954 Drosophila
C 27	46.4	6.4	796	9	CG271311	CG1BR78TV
C 28	46.2	6.3	564	6	CB815935	3529.1.78
C 29	46.2	6.3	813	9	CG450503	CG450503
C 30	46.2	6.3	918	9	CG643169	CG643169
C 31	46	6.3	517	2	BE400971	BE400971
C 32	46	6.3	592	7	CF132851	CF132851
C 33	46	6.3	1009	9	CNS0108W	AL098882 Drosophila
C 34	45.8	6.3	437	4	B1679572	B1679572
C 35	45.8	6.3	524	6	CD485064	CD485064
C 36	45.8	6.3	627	7	COS33576	COS33576
C 37	45.8	6.3	763	7	COS26309	COS26309
C 38	45.6	6.3	1366	8	BZ572253	BZ572253
C 39	45.6	6.3	2082	7	CL980714	CL980714
C 40	45.4	6.2	560	7	CN887421	CN887421
C 41	45.4	6.2	574	7	CN898104	CN898104
C 42	45.4	6.2	695	9	CL541678	CL541678
C 43	45.4	6.2	1005	9	CL974026	CL974026
C 44	45.2	6.2	726	9	CL963016	CL963016
C 45	45.2	6.2	811	7	CN152570	CN152570

ALIGNMENTS

RESULT 1
LOCUS BG846752 666 bp mRNA linear EST 29-MAY-2001
DEFINITION 1024014H04.Y1 C. reinhardtii CC-1690, normalized, lambda zap II
ACCESSION BG846752
VERSION BG846752.1 GI:14227936
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE 1 (bases 1 to 666)
Grossman, A., Davies, J., Pederspiel, N., Harris, E., Lefebvre, P.,
McDermott, J. P., Sillflow, C., Stern, D. and Surzycki, R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants: Project phase 2
Unpublished (2000)

JOURNAL
COMMENT
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source
Location/Qualifiers

1..666
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_1ib="C. reinhardtii CC-1690, normalized, lambda zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI, Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown in the mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldi et al (1996) Genome Research 6: 791-806."

Db 1178 SGYKTTSSCCSCCKGKSTCKGCKKCKSGCKGBKGBGAGCCBCCCTTSSCKCCCKBKSCKGCKG 111 191
 Oy 84 CCTGATATGCAATGGCGTGTGTGGTTTCTGTGAGACGGCGCGCGCATCCCATCTCGCGGTGCG 143
 Db 1118 CTGTBCGCCCCSCTCSGCCGSSGCGCTGSSGCGCGCGCGSSGSSSCTCYGCGKGGCGSC 10559
 Oy 144 GAATTTCTGTGGGGAGCTGACCTGAGTGTGCGATCGATGTGTTTCAATCAGCGGCATGACGCAT 203
 Db 1058 GKGSGBGCTGAGKGBTKKGGSKCCCBKBSCKGCKRYGTTTGCKKKKBSYTKTCCCKCY 999
 Oy 204 GCATGAGGTGAGTGTGTGCGCGGGAGCGCCGCGCGCCCAATGCGGCGAGTGGCCAGCTT---GT 260
 Db 998 KGBKTKTCCCTCCCCCCTTGGSSYSGGCGCCSCCCCCCCCCCTTBGSHBTKSKGTGKGS 939
 Oy 261 CCTGTGCGTGTATGCGCGGATTTTCTGTGGCGCAAGATGATCGTCAAGCACAATGCGCCATCA 320
 Db 938 CCGKCGCTGTGTBYTCTGTGGGCTGTBKTKGKGGGGKKSCKMGATKTTGCTBBSKCSGYCC 879
 Oy 321 TCGCCATGCGGGAACCGACGACGCCA-----GATTTCAGCATGCGGGCCCGGTGCG 374
 Db 878 BCATGCGCMCKGTTTCCSCCDPTCCCCCCCCCGGSCCTTTCTYCTTCCKGBCBMCCM 819
 Oy 375 CTGGTACGCCCGCGCTTCAATCGGACCTAATTTGCGCTGCGCGAGGGGCGTGTGCTGCCGT 434
 Db 818 CYKGCBCGSGYKRYTKKSKSGCMCMVMGVTKGKKGKGGGGGASGGAATATMASAGSMRAM 759
 Oy 435 CATGTGACGAGTCTATGCGCTGATGTTGGGGAGATCGCTGATGTGAGTG-----TCTTC 489
 Db 758 AGMAAGMAMMAGMAAKRAGAGGTGKGGGRAMRAMRAYAANGAKAKRMMCMCTMGGG 699
 Oy 490 TGGCCGTTGCCGTGATCCTGCGCGTCATCCAGCTGTTCGTGTTCCGAGATCTGAGTGGCG 549
 Db 698 KGGGGGKKAACMGGGSGVMGVSQVMCVMMGCGCGSGCGGGGKSGSGCMNMNGKGTMG 639
 Oy 550 CACGCGCCCGGACACGACGCGTCCCGGACCGCCACAATGCGCGGCTGTCGCGATCAGC 609
 Db 638 SCCGCGCCCGCKCMGACCMGCTCGCGGSGYCGCGCCCTGCBBCBTKBTBTTCGCTCKC 579
 Oy 610 GACCCCGGTGCGTGTGACCTGCTTTCACCTT 642
 Db 578 GTNSCCTGTCCTCYBTBTCTCTCTTCTT 546

 RESULT 4
 CNS0175/c
 LOCUS
 DEFINITION
 Drosophila melanogaster genome survey sequence SP6 end of BAC
 BACN3J7U08 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION
 AL108460
 VERSION
 AL108460.1 GI:5628764
 KEYWORDS
 GSS.
 ORGANISM
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 REFERENCE
 Genoscope.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 COMMENT
 - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.ebi.ac.uk-. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBeloBAC11.
 FEATURES
 Location/Qualifiers
 1..1101

Query Match 6.9%; Score 50.6; DB 9; Length 1101;
Best Local Similarity 13.4%; Pred. No. 0.17;
Matches 42; Conservative 158; Mismatches 113; Indels 0; Gaps 0;

ORIGIN

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37L08"
/clone_id="DrosBAC"
/plasmid="pBelosBAC11"
/note="end = SP6"

Dy 322 CGCATGCCGAACCCAGCAGACCAGATTTCACATGCGGGCCCCGCTCCGTGCATG 381
::
Db 1069 SBTSTTSTTBTSBSSSSBSBSBSGSSBSBSSTSSSSSSTSSSSTSBTBRTS 1010A
::
Dy 382 GCCGCTTCATCCGCACTTAATTGCGTGCGCAGAGGGGCTGCTGCCGTCATCGTG 441
::
Db 1009 TSSSTTTTTTTTTTTTSTBTBBSBSBSBSSTSSSBRTSKBTBSBSBSTSSASBS 950
::
Dy 442 ACGGTATAGCGGTATGTGGGGATCGCTGATGTACGTGTCTTGCGCCCTGGC 501
::
Db 949 SSSSSBSSTSTBTSTBBBSSTSSSGSSSSSBSTBSBSBRTTTTBTKSTSTSS 890
::
Dy 502 TCATCTCGGCGTCGATCCACTCTTCTGTTCGGCATCTGCTGCCGACCCGCCGC 561
::
Db 889 STSTBTSTBSHSSSSSSBSBSSTTBSBSSTSYSSSBSTCKSTBTBSGTBTB 830
::
Dy 562 CACAGACGCTCCCGACCGCACATATGCGGCTGCGGATCAGACCCCGTCGTC 621
::
Db 829 MSKBSSTSTSTBSBSBGGSBGSSCGSTGSSCSBCHGSTSSGTSSTTGCBCTGGCS 770
::
Dy 622 CTGCTGACTGCT 634
| : : | : | : |
Db 769 CCCGCCCTCCT 757
| : : | : | : |

RESULT 5
CNS0091P/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CNS0091P 925 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TE3 end of BAC #
BACR19D16 of RPI1-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL053013
AL053013.1 GI:4934461
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Epiphytoidae; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ::
BP 191 91006 EVRY cedex - FR:NCR (E-mail: seqref@genoscope.cns.fr
- Web: www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mosmoser in Peter de Jong's laboratory in the department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the library
PI and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.


```

/clone="TA15020A09"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/clone_id="TA1515"
/notes="vector: pCMV-sPORT6.0 (Invitrogen Technologies);
Site_1: Nc1; Site_2: Mui; mRNA obtained from wheat seeds
of cultivar Glena 15 days post-anthesis"

```

Query Match	6.7%;	Score 48.8;	DB 5;	Length 427;
Best Local Similarity	46.3%;	Pred. No. 0.44;		
Matches 161;	Conservative	0;	Mismatches 187;	Indels 0;
				Gaps 0;

0y 214 GTGCTGCGCGGGGCGCGCGCGCCAAATGGCGGCGAATGGGCGAGCTTGCTGTGGCTGAT 273
 Db 59 GTCCCGAGCAGGGCCCCCGCGCGCGCGCGCGGTGCTCTCTTCACGGAGCTTCCGGAG 118

Qy 274 GCCGATTTTCCCTGGCGCAAGATATATCGTCAAGACATGGCCCATCATGGCCCATGGCGGA 333
Db 119 CTCTGGCTTCTCTGGCGCCACCAAGATGGCGCGCTTTCGCGCGCCCGCGGCGCTTTCGCGCGCCCTC 178

Qy	334	ACCGAGACAGACCCGAGTTTGGACATGGCGGCCCGTCCGTGTAGTACGCCCGCTTCATC	393
Db	179	GCCCCGACCTCCGCGCTACGGGCACTCCGACGGCCCCCGCGGACCCCGCGCTACACC	238

QY 394 GGACACCTATTTCGGCTGGCGCGGAGGGGCTGCTGCTGCCCTCATGTATACGGTATACGG 453

Db 239 ATGCTCCACGTCGTGCGCGACGTAAGTGGCTCTCTCGACACCACTTCGCGCTCCCAAGG 298

QY 454 CTGATGTTGGGGGATCGCTGGATGTACGTCGTCTTCTGGCCGTTGCCGTCGATCTTGAGC 513

Db 239 GTGTGTGGGGTCAAGCTGGGGTGGCGCAGGTGGCGTGCATTTCTGCTGTTCGGCGC 358

514 TCGATCCAGCTGTTTCGTGTCGATCTGGCTGCGACCGCCCCCGC 561
 359 GACCGGGTCCGCGCCGTTGCTGCGCTGGGATCCCGTTTACCCGCG 406

RESULT 8	CA108291	LOCUS	DEFINITION
CA108291	663 bp	mRNA	linear
SCSFHR1043A02.9	HR1	Saccharum officinarum	CDNA clone SCSFHR1043A02
5', mRNA sequence.			

VERSION	CA108291.1	GI:34961598
KEYWORDS	Est.	
SOURCE	Saccharum officinarum	
ORGANISM	Saccharum officinarum	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 663)
Vettore, A.L., da Silva, F.R., Kemper, E.L., and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
complex.
clade); Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD

Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parreira@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
<http://www.bcccenter.fcav.unesp.br>
Plate: 043 row: A column: 02
Seq primer: T7 Promoter Primer.

FEATURES	location/qualifiers
source	1..663
	/organism="Saccharum officinarum"
	/mol_type="mRNA"

```

/zb xref="taxon:4547"
/clone="SCSFHR1043A02"
/lab_host="DH10B"
/clone_1lb="HR1"
/notes="Organ: seedlings inoculated with Herbaspirillum
rubi/subalbicans; Vector: pSPort; Site_1: Salt; Site_2:
Nori; An unidirectional cDNA library generated from
[seedlings inoculated with Herbaspirillum
rubi/subalbicans]. cDNA was prepared from poly(A+ mRNA
using Superscript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucsec.lad.ic.unicamp.br/public"

```

Query Match	6.7%	Score 48.8;	DB 6;	Length 663;
Best Local Similarity	45.8%;	Pred. NO. 0.44;		
Matches 167;	Conservative	0;	Mismatches 199;	Indels 0;
				Gaps 0;

Oy 214 GTCGTGCGGGGCGCCCCGCGCCGACATGCGCGCATGGGCGACGTTGTCCTGAGCTGAT 273
 Db 43 GTTGC CGGACTGACTTCGGCGCGTCCCGCGTCCCGCGCCCTTCCTGCTCTCTCTGCG 102

Oy 274 GCCGATTTTCTGGCGCGAAGATGATGTCGTCAGACATGGCCCATCATGCGCCATGCGGA 333
 Db 103 CTCCTCGGTTCCGCGCCCTCGCGATGGAGGAGCAGCATGGCCGCGCACGTCCTCTGGCGGC 162

QY 334 ACCGACGACGACCCAGATTTCGACCATGCGCGCCGGTCGGTCGTGACGCCCCCTTCATC 393
 DP 163 GTGAAGGACCAATCCCGCGCGCGCCGACGCGCGGATTCGACGCGGGCTCGGCGCTTCGCC 222

394 GGCACCTATTTCGCGTGGCGCGAGGGGCTGCTGCTGCTGCCGTGATGAGACGGTCTATGCG 453
 454
 223 GTTGCACGACCAACACAGGCGCGGAGACGGCTGTGAGATTGCTGCGCGGTGTGAGAGGCC 282

454 CTGATGTTGGGGGATCGCTGAGTGTACGTGTCCTTCTTG6GCGCTGTGCGCTGATCCTGGCG 513

Oy 514 TCGATCCAGCTGTTGTGTTCGGCATCTGGTGCACGCCGCCGCCGACGACGAGCTTC 573
Db 343 ATGTAAAGACTCTTAACCAAGCCCAAGCATTGGTGTTAAGTAAATGCTTGAATTCATGAGAGCTC 402

QY	574	CCGA	578
	1	111	
AO3	0.0001		AO7

RESULT 9			
CNS016KD	LOCUS	DEFINITION	
CNS016KD	1100 bp	DNA	linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence Sp6 end of BAC			
BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit			
fly), genomic survey sequence.			
AJ106A55			

VERSION	AL106855.1	GI:5624152
KEYWORDS	GSS.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	
	Phylum: Arthropoda	
	Class: Insecta	
	Order: Diptera	

ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydraoidea; Drosophilidae; Drosophila.
1 (bases 1 to 1100)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Séquençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqret@genoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC

Query Match	Similarity	Score	48:	DB	7:	Length	613:
Best Local	Similarity	44.5%	Pred	No	0.68:		
Matches	241:	Conservative	0:	Mismatches	295:	Indels	6:
						Gaps	1

Query	35	CGCCACCAAGTTTGATCGTCTCGGAGGAGCATATCGCCGCGTGGCCCTGACATGTGC	94
Db	29	CGCCGATGCTTGGTGTGGAGGCTGCTCGGCGGTGCGCGCGTACCGAGCATTCGGCGGA	88
Qy	95	ATGCGCTGTGGTTCTTGACACGGGCGGCGATCCATCTCGGCGGTTCGGAAATTTCCTG	154
Db	89	ACAGATCTCTTGGCAAGGACGGGCGGCGCGGCATCTGTGCTCTCGAAGACAGACGG	148
Qy	155	GAGTACCTGAGCTGTGCGGTGCTGTGTTTCATCATCGCGATACGACGATGACATGAGTGG	214
Db	149	TCCGCGCCGCGCGCGCGCTGCGCGGCGGTGCGAGTGAAGAACTCAGGCTCTCGGTGTG	208
Qy	215	TCGTGCGCGAGGCGCCGCGCGCAATGCGGCGATGAGGCGCACTTGTCTGTGGCTGTATG	274
Db	209	ACATGCCGATGTATGTGTCACATACATACAGAGGGGCTGTCTTCGCGCGCTCTTC	268
Qy	275	CCGATTTTCTCTGGCGCAAGATGATGTCACACATGAGCCATATGCCATGCGCGAA	334
Db	269	CCGGGCTCTGGCGGCTCTGCTGTGTGTGCTGCTGCGGCGGCTCTCCCGCGCTCG	328
Qy	335	CCGACGAGGACCCCAATTTTCAGCAGATGAGCGCGCGGCTCGGTATAGCGCGCTTCATG	394
Db	329	GGTTTCTCCGCGCTCGCGCGCGCTGCTGACGCTC-----ACGACGACGATCG	382
Qy	395	GCACTTATTTTCGCTGTGCGCGGAGGAGCTGTGCTGCTCCGCTCATCTGTGACGTCTAATG	454
Db	383	TGATACGGTGTGCGCGCGCGACAGACGCGCGCTGTGATTTACACACGCGGTGTGCGCTC	442
Qy	445	TGATTTGGGGAGATGCTGTGATGTACGTGTCTTTCGCGCTGCGCTGTGATCTGCGCT	514
Db	443	TGTCCCTCGGCGCACTTCTGTGTGCGCGCGCGCGACGCTGCGACCGGTTGAGAAAGCGC	502
Qy	515	CGATCCACCTGTGCTGTGTTGCGCATCTGCGCTCGGACCGCGCGCGGACGAGCGGCTTC	574
Db	503	TGTTTCCCATGAGACCGGACGCTGACGCTACGCGCGCGACGCGCGCTCACGGGCTTC	562
Qy	575	CG 576	
Db	563	AG 564	

RESULT 13	CG236345	785 bp	DNA	linear	SSS 22-AUG-2003
LOCUS	CG236345				
DEFINITION	OGMLC52TV_ZM_0.7.1.5_KB_Zea mays genomic clone ZMMBMA0667107,				
ACCESSION	CG236345				
VERSION	CG236345.1				
KEYWORDS	SSS.				
SOURCE					
ORGANISM	Zea mays				
REFERENCE	Zea mays				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoidae; Andropogoneae; Zea.				
	1 (bases 1 to 785)				
	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Renick, A., Frazer, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.				
	Consortium for Maize Genomics				
	Unpublished (2002)				
	Other SSSs: OGMLC52TH				
COMMENT	Contact: Cathy Whitelaw				
	TIGR				
	9712 Medical Center Drive, Rockville, MD 20850, USA				
	Tel: 301-838-5843				
	Fax: 301-838-0208				
	Email: whitelaw@cigr.org				
	Seq primer: TP				
	Class: sheared ends.				

FEATURES	source	location/Qualifiers
ORIGIN		1..785 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBMA0667107" /clone_1lb="ZM 0.7 1.5 KB" /note="Vector: pBCSK-, Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
Query Match	6.6%; Score 48; DB 9; Length 785;	
Best Local Similarity	44.5%; Pred. No. 0.69;	
Matches 241; Conservative	0; Mismatches 225; Indels 6; Gaps 1;	
QY	35	CCGCCACCAAGTTGATGCTCTCGGGCGGCATCATGCGCCGCTGCTGACCTCTGATGTC 94
DB	209	CCGCCAGATGCTTGTGTGAGAGCTCTCTCGGCCGCGCGCGTACCAGACATCGCGCGA 268
QY	95	ATGCGCGTGTGTTCTTGAGACGGCGCGCGCATCCCATCTCGCGGTGGCCAATTTCTCGG 154
DB	269	ACAGAAATCTTTTGCMAAGACGCGCGCGCGCATCTGTGTCGAAAGCAGACGG 328
QY	155	GGCTGACCTGGCTGTGCGTGTGCTGTTCATCATCGGCAGCAGACGCGATGATGGTGGG 214
DB	329	TCCGACCCCGCGCGCGCTGTGCGCGCGGTGCGCACTGAGACCTTAGCTCTCGGTGTGCG 388
QY	215	TCTGTCGGGGCGCGCCGCGCCAAATCGCGGATGGAGCGCAAGTGTCTGTGCTGTATG 274
DB	389	ACATGCGGATGTCTATCGTGCCACTACATCCAGAAAGGGGCTGTCTTCCGCGCGCTCTC 448
QY	275	CCGATTTTCTTGCGCGCAATGATGCTCAAGCAGCATGCGCCATATCGCCATGCGCATG 334
DB	449	CCGCGCTGCTGCGCGCGCTGTGCTGTGCTGTGCTGTGCGCGCGCTCTCCGCGCGCTCG 508
QY	335	CCGACGACGACCCAGATTTTGACCATGAGCGCGCGCGCTGCTGTGATGCGCCGCTTCATCG 394
DB	509	GCGTCTTCCCGCTCGCGCGCGCGCGCTCTGTCAAGCTCTC-----ACGACGACGATCG 562
QY	395	GCACTTATTTTTCGCTGTGCGCGCGAGGGGCTGTGCTGCCCGTATCATGATGATCGTCTATGCG 454
DB	563	TGATACGGTGTGCGCGGAGAGACGCGCGCGCTGTCTACACAGCGGATGGCGCGCTCTC 622
QY	455	TGATGTTGGGGATGCTGATGTAAGTGTCTTTCTGCGCGTGTGCGCTGATCTGCGCTG 514
DB	623	TGTGCTGTGCGGATCTTCTGTGCTGCCGCGCGCGCATGTCGCGACCCGTTTGACGAACGCGC 682
QY	515	CGATTCGACGCTTGTGTGTGCGCATCTGCGCTCGCGCGCACCGCGCGCGGACAGCGCGTTCG 574
DB	683	TGTTTCCCATGACCGGACGGTGAAGTACAGCGCGCGCGCGCGCTCAGCGCGCTTC 742
QY	575	CG 576
DB	743	AG 744
RESULT 14		
CA096815	634 bp	mrna linear EST 23-SEP-2003
LOCUS	SCCCCL6005G08.g	CL6 Saccharum officinarum cDNA clone SCCCL6005G08
DEFINITION	5', mRNA sequence.	
ACCESSION	CA096815	
VERSION	CA096815.1	GI:34950122
KEYWORDS	EST.	
SOURCE	Saccharum officinarum	
ORGANISM	Saccharum officinarum	
REFERENCE	1 (bases 1 to 634)	
AUTHORS	Vertore,A.L., da Silva,F.R. Kemper,E.L. and Arruda,P.	

TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parnuda@unicamp.br

FEATURES
 source
 1..634
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCCCL6005G08"
 /lab_host="DH10B"
 /clone_lib="CL6"
 /note="Organ: Pool of sugarcane calli submitted to low (40C) and high (37 C) temperature stress; Vector: pSport1; Site 1: Sali; Site 2: NoTi; An unidirectional cDNA library generated from [Pool of sugarcane calli submitted to low (40C) and high (37 C) temperature stress]. cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucet.1ad.ic.unicamp.br/public>"

ORIGIN

Query Match 6.5%; Score 47.6; DB 6; Length 634;
 Best Local Similarity 48.0%; Pred. No. 0.85;
 Matches 171; Conservative 0; Mismatches 179; Indels 6; Gaps 1;

QY 36 GCGCACCAAGTTGATGTCCTCGGCGGCATCATCGCCGCTGCTGCGCCCTGCATGTGA 95
 DB 260 CCGCACCAAGTTGATGTCCTCGGCGGCATCATCGCCGCTGCTGCGCCCTGCATGTGA 319
 QY 96 TCGCGTGTGTTTCTGAGCGCGCGCGCATCCATCCGTGGCGGTGCGCAATTTCTGGG 155
 DB 320 TCTGCAAGTGTCTCGACCGGCTGACACAAAGCGCGGTCTGCTGCAATGTTCTTTGA 379
 QY 156 GCTGACCTGCTGTGCGTGTCTGTTTCATCATCGGCATGACGCGATGCATGGTGGT 215
 DB 380 TCTAAACAGTGTACGATGACGAGTCTATCGCCGCGCTGCGCTTTGAGAGAACGCG 438
 QY 216 CGTGCGGCGCGCGCGCGCGCATGCGCGATGGCGAGCTTGTCTGTGCGTGTATGC 275
 DB 439 -----CGAAGCCCAAGAGAGTCAACAGCCATGGCGCGCGCTGATCTGTGCAAGAGA 493
 QY 276 CGGATTTTCTGCGCGCAAGATGATGTCAGACCATGCGCCATCATCGCCATCCCGGAG 335
 DB 494 CTGGAGGAGCATCTCGAAGCGCGGTGCGACGAGAGACTTCAAGCGGAGTGGCGGCTT 553
 QY 336 CGACGACGACCAAGATTTTGACCATGCGCGCGCGGTGCGCTGTTAGCGCCGCTTCA 391
 DB 554 CAGCAACCAAGCGCAAGCGCGCGCGCGCGCTTGGCGCTTGGCGCGCATTTCA 609

RESULT 15 BG343324 933 bp mRNA linear EST 22-OCT-2001
LOCUS BG343324
DEFINITION HVSMEg0005G17f Hordeum vulgare pre-anthesis spike EST library
 HVCNDA0008 (white to yellow anther) Hordeum vulgare subsp. vulgare
 cDNA clone HVSMEg0005G17f, mRNA sequence.
ACCESSION BG343324
VERSION BG343324.1 GI:13155653

KEYWORDS
SOURCE EST.
ORGANISM Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

REFERENCE
 1 (bases 1 to 933)
 Wing,R., Close,T.J., Kleinbofs,A., Wise,R., Begum,D., Fritsch,D.,
 Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,
 Denton,R.D., Close,S.J., Oates,R. and Main,D.
 Development of a genetically and physically anchored EST resource
 for barley genomics: Morex pre-anthesis spike cDNA library
 Unpublished (2001)
JOURNAL
 CONTACT: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 474
 Seq primer: AATTACCTCACTAAGCG
 High quality sequence stop: 744.
COMMENT
 Location/Qualifiers

FEATURES
 source
 1..933
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Morex"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="HVSMEg0005G17f"
 /tissue_type="pre-anthesis spike"
 /lab_host="SOLR"
 /clone_lib="Hordeum vulgare pre-anthesis spike EST library
 HVCNDA0008 (white to yellow anther)"

/note="Vector: lambdaZAP; Site 1: Scoli; Site 2: Xho1;
 Plants were grown in the greenhouse at the University of
 California, Riverside (Penton, SJ Close, TJ Close). Whole
 spike with awns trimmed were collected at white, green and
 yellow anther stages (Fenton). Total RNA was prepared from
 each pool, equal quantities of all three RNA pools were
 combined, poly(A) RNA was purified from the mixture, one
 primary unamplified cDNA library was made, and 1 million
 pfu were in vivo excised to give plasmid SK(-) cDNA
 phagemids. These steps were performed in the TJ Close lab
 (Choi) at the University of California, Riverside.
 Phagemids were plated and picked at the Clemson University
 Genomics Institute (CUGI) (Begum, Palmer, Fritsch, Atkins
 and Wing) plasmid DNA preparations, DNA sequencing and
 sequence analysis were performed at CUGI (Wing, Yu,
 Fritsch, Henry, Simmons, Oates, Rambo, Main). The sequence
 has been trimmed to remove vector sequence and contains a
 minimum of 100 bases of phred value 20 or above. For more
 details on library preparation and sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
 this clone see <http://www.genome.clemson.edu/orders> Also
 see Close TJ, Wing R, Kleinbofs A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (<http://wheat.pw.usda.gov/g99pages/dgn/31/cover.html>)"

ORIGIN

Query Match 6.5%; Score 47.2; DB 4; Length 933;
 Best Local Similarity 46.0%; Pred. No. 1.1;
 Matches 160; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 214 GTCTGCGCGCGCGCGCGCGCAATGCGCGGATGCGGCACCTGTCTGTGCTGTAT 273
 DB 80 GTCTGCG 139
 QY 274 GCGGATTTTCTGCGCGCAAGATGATGTCAGACCATGCGCCATCATCGCGATGCCGGA 333
 DB 140 CTCTGCGCTCTCTGCGCGCAAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 199

```
Oy 334 ACCGACGACGACCAGATTTCAGCCATGCGGCCCGCTCCGCTGGTACGCCCGCTTCATC 393
    |||||
Db 200 GCGCCCGACCTCCGGGCTACGGCGACTCCGACGCCCGGACAGACCCCGCCCTACACC 259
    |||||
Oy 394 GGCACCTATTTCGCTGCGCGAGGGGCTGCTGCTGCCCGTATCGTAGAGGTCTATGCG 453
    |||||
Db 260 ATGCTTCAGTGTGCGGACGCTCGTCCGCTCTCGACCACTCCGCTCCCAAGGTG 319
    |||||
Oy 454 CTGATCTTGGGGGATCGCTGATGTACGTGTCTTCTGSCCGTTCGCCGTGATCCTGAGG 513
    |||||
Db 320 CTGTGCGGGGACGACTGGGGCGGCGAGTGGCGTGGACCTCTGCTGTTCGGGCG 379
    |||||
Oy 514 TCGATCAGCTGTCTGTTCGGGCACTGGCTGCCGACCGCCCGGC 561
    |||||
Db 380 GACCGGTTCCGGCGCTCGCTGGGATCCGTTCTTCCCGGC 427
    |||||
```

Search completed: June 22, 2005, 01:01:23
Job time : 2877.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 18:02:15 ; Search time 3377.93 Seconds
(without alignments)
10457.239 Million cell updates/sec

Title: US-10-695-980-31
Perfect score: 729

Sequence: 1 atgagcgccatcgcctcgtcc.....ccaagggggacacgcgcatga 729

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_scs: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	729	100.0	729	6	AB4691 Sequence 1
2	729	100.0	729	6	AR452029 Sequence 1
3	729	100.0	729	6	AX701576 Sequence 1
4	729	100.0	1261	6	AR169854 Sequence 1
5	729	100.0	1261	6	AR452028 Sequence 1
6	698.6	95.8	729	6	CQ788339 Sequence 1
7	698.6	95.8	729	6	CQ793306 Sequence 1
8	698.6	95.8	729	6	CQ793487 Sequence 1
9	698.6	95.8	6335	1	PMCRWXYI
10	697	95.6	729	6	CQ788335 Sequence 1
11	697	95.6	729	6	CQ788369 Sequence 1
12	697	95.6	729	6	CQ793302 Sequence 1
13	697	95.6	729	6	CQ793396 Sequence 1
14	697	95.6	729	6	CQ793483 Sequence 1
15	697	95.6	729	6	CQ793561 Sequence 1
16	697	95.6	2886	6	AR041164 Sequence 1
17	697	95.6	2886	6	AR082018 Sequence 1
18	697	95.6	2886	6	AX429767 Sequence 1
19	697	95.6	5373	1	ATUCRTWA

20	611.8	83.9	639	6	AR041161 Sequence 1
21	611.8	83.9	639	6	AR082015 Sequence 1
22	611.8	83.9	639	6	AX429761 Sequence 1
23	446	61.2	1002	12	AY345165 Sequence 1
24	446	61.2	5397	12	AY605097 Sequence 1
25	434.2	59.6	902	6	AX350856 Sequence 1
26	434.2	59.6	902	6	BD073961 Sequence 1
27	432.4	59.3	729	6	AR041165 Sequence 1
28	432.4	59.3	729	6	AR082019 Sequence 1
29	432.4	59.3	729	6	AX429768 Sequence 1
30	432.4	59.3	1631	1	AFACRTWB
31	432.4	59.3	1631	6	AR041167 Sequence 1
32	432.4	59.3	1631	6	AR082021 Sequence 1
33	432.4	59.3	1631	6	CQ788337 Sequence 1
34	432.4	59.3	1631	6	CQ793304 Sequence 1
35	432.4	59.3	1631	6	CQ793485 Sequence 1
36	432.4	59.3	1631	6	AX429772 Sequence 1
37	383.8	52.6	726	6	A58570
38	383.8	52.6	726	6	A84707
39	383.8	52.6	726	6	AR103084 Sequence 1
40	383.8	52.6	726	6	AR139700 Sequence 1
41	383.8	52.6	726	6	AR169853 Sequence 1
42	383.8	52.6	726	6	AR391879 Sequence 1
43	383.8	52.6	726	6	AR452027 Sequence 1
44	217	29.8	776	6	CQ788343 Sequence 1
45	217	29.8	776	6	CQ793310 Sequence 1

ALIGNMENTS

RESULT 1
LOCUS AB4691
DEFINITION Sequence 1 from Patent EP0872554.
ACCESSION AB4691
VERSION AB4691.1 GI:6733567
KEYWORDS
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 729)
AUTHORS Pasamontes, L. and Teygankov, Y. S.
TITLE Improved fermentative carotenoid production
JOURNAL Patent: EP 0872554-A1 21-OCT-1999;
HOFFMANN LA ROCHE (CH)
FEATURES
source location/Qualifiers
1..729
/organism="unclassified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 729; DB 6; Length 729;
Best Local Similarity 100.0%; Pred. No. 3.3e-110;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGCATGCGCTGCCAAGCGAGATGACCGCCACGATTGATGCTTCGGGC 60
1 ATGAGCGCATGCGCTGCCAAGCGAGATGACCGCCACGATTGATGCTTCGGGC 60
1 ATGAGCGCATGCGCTGCCAAGCGAGATGACCGCCACGATTGATGCTTCGGGC 60
QY 61 GGCATCATGCGCGCGTGGCTGCGCTGCAATGCGCGCTGCTGCGCTGCGCGG 120
61 GGCATCATGCGCGCGTGGCTGCGCTGCAATGCGCGCTGCTGCGCTGCGCGG 120
61 GGCATCATGCGCGCGTGGCTGCGCTGCAATGCGCGCTGCTGCGCTGCGCGG 120
QY 121 GGCATCATGCGCGCGTGGCTGCGCTGCAATGCGCGCTGCTGCGCTGCGCGG 180
121 GGCATCATGCGCGCGTGGCTGCGCTGCAATGCGCGCTGCTGCGCTGCGCGG 180
121 GGCATCATGCGCGCGTGGCTGCGCTGCAATGCGCGCTGCTGCGCTGCGCGG 180
QY 181 TTTCATCATGCGCGATGCGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 240
181 TTTCATCATGCGCGATGCGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 240
181 TTTCATCATGCGCGATGCGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 240

QY	241	GGGGGAGATGGGGCACACTGTTGCTCGTGGCGCTGATGCGGATTTTCCGGGGGCAAGATGAAC	300
Db	241	GCGGCGATGGGCGACAGTTGCTCTGGCGCTGATGCGGATTTTCCGGGGGCAAGATGAAC	300
QY	301	GTCAAGACATATGGCCATTCATGCGCATGCGGAAACCGACGACGACCCAGATTTTCAGCAT	360
Db	301	GTCAAGACATATGGCCATTCATGCGCATGCGGAAACCGACGACGACCCAGATTTTCAGCAT	360
QY	361	GGCGGCGCGGATCCGCTGGTAGGCGCGGCTTCATGCGACCTATTTGGGCGTGGGCGACGGGG	420
Db	361	GGCGGCGCGGATCCGCTGGTAGGCGCGGCTTCATGCGACCTATTTGGGCGTGGGCGACGGGG	420
QY	421	CTGCTGCTGCCCCGTCATGCTGACGGCTCTATGCGCTGATTTGGGGAGATGCTGAGATGAC	480
Db	421	CTGCTGCTGCCCCGTCATGCTGACGGCTCTATGCGCTGATTTGGGGAGATGCTGAGATGAC	480
QY	481	GTGGCTTTCTGGGCGGTTGGCCGTCGATCCGCGGGGTGATNCCAGACTGTTGCGTTTGGGATAC	540
Db	481	GTGGCTTTCTGGGCGGTTGGCCGTCGATCCGCGGGGTGATNCCAGACTGTTGCGTTTGGGATAC	540
QY	541	TGGCTGCGGACCGGCGCCGCGGACGACGCGGTTCCGCGACCGGACCAATATCGCGGTCGTG	600
Db	541	TGGCTGCGGACCGGCGCCGCGGACGACGCGGTTCCGCGACCGGACCAATATCGCGGTCGTG	600
QY	601	CGGATCAGCGAACCCTGTGCTGCTGACCTGCTTCACTTTGGCGGTTATCATACGAA	660
Db	601	CGGATCAGCGAACCCTGTGCTGCTGACCTGCTTCACTTTGGCGGTTATCATACGAA	660
QY	661	CACCACTGACCCGACCGGTGCTTGGTGGCGGCTTGGCCAGACACCGGACCAAGAGGGGAGC	720
Db	661	CACCACTGACCCGACCGGTGCTTGGTGGCGGCTTGGCCAGACACCGGACCAAGAGGGGAGC	720
QY	721	ACCGATGA 729	
Db	721	ACCGATGA 729	

RESULT 2
ARA52029

DEFINITION	Sequence 31 from patent US 6677134.
ACCESSION	AR452029
VERSION	AR452029.1
KEYWORDS	GI:42683356
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 729)
TITLE	Paramontes, L. and Tsygankov, Y.
JOURNAL	Fermentative carotenoid production
FEATURES	Patent: US 6677134 1991-13-94N-2004; Location/Qualifiers

```

/organism="unknown"
/mol_type="genomic DNA"

```

Query Match	100.0%	Score 729	DB 6	Length 729	
Best Local Similarity	100.0%	Pred. No. 3.3e-110			
Matches 729	0	Mismatches	0	Gaps	0

QY	DB
1 ATGAGCGCATGCGCTGACCAAGCAGACTGACCGCACCCAGTTGATGTCGCGGC 60	1 ATGAGCGCATGCGCTGACCAAGCAGACTGACCGCACCCAGTTGATGTCGCGGC 60

QY 61 GGATCATCGCCGCGCTGGCGCTGGCATGTGCATGCGCTGTGGTTTCTGGAGCGGCG 120

Dp 61 GGATCATCGCCGCGCTGGCGCTGGCATGTGCATGCGCTGTGGTTTCTGGAGCGGCG 120

QY 121 GCGCATCCCATCTCTGCGCGTGCAGAAATTTCTTGCGGCTGACCTGCTGTCTGCTGCGGTCTTG 180

Db 121 GCGCATCCCATCTCTGCGCGTGCAGAAATTTCTTGCGGCTGACCTGCTGTCTGCTGCGGTCTTG 180

OY	181	TTCAATCATCGGCAATGAAGCCATGATGAGGTCCGGTCGCGCGGGGCGCCGCGCCCAAT	240
Db	181	TTCAATCATCGGCAATGAAGCCATGATGAGGTCCGGTCGCGCGGGGCGCCGCGCCCAAT	240
OY	241	GGGGCGATGGGCAAGCTGTCTCTGTGCTGTATGCGGATTTTCTTGGCGAAGATGATC	300
Db	241	GGGGGAGATGGGCAAGCTTGTCTCTGTGTGTATGCGGATTTTCTTGGCGAAGATGATC	300
OY	301	GTCAGACATATGGCCCATATCGCCATATCCGGAAACGACGACGACCCGATTTGCACCAT	360
Db	301	GTCAGACCATATGGCCCATATCGCCATATCCGGAAACGACGACCCGATTTGCACCAT	360
OY	361	GGGGCCCCGGTCCGCTGTGTACGCCCGTTCAACGGCACTAATTTTGGCTGGGCGGAGGGG	420
Db	361	GGGGCCCCGGTCCGCTGTGTACGCCCGTTCAACGGCACTAATTTTGGCTGGGCGGAGGGG	420
OY	421	CTGCTGTGACCCCGTCATCGTAGACGGTCTATGACGCTGATGTGGGGGATCGCTGATGTATC	480
Db	421	CTGCTGTGACCCCGTCATCGTAGACGGTCTATGACGCTGATGTGGGGGATCGCTGATGTATC	480
OY	481	GTGGTCTTTCTGGCCGTTGCGGTGCATCTGTGGCGTGCATCCAGCTGTTCTGTTCGGCATC	540
Db	481	GTGGTCTTTCTGGCCGTTGCGGTGCATCTGTGGCGTGCATCCAGCTGTTCTGTTCGGCATC	540
OY	541	TGGCTGCCCGACCGCCCGCGGCAAGACGCGCTTCCCGGACCGCCCAAAAGCGCGGTCTGCG	600
Db	541	TGGCTGCCCGACCGCCCGCGGCAAGACGCGCTTCCCGGACCGCCCAAAAGCGCGGTCTGCG	600
OY	601	CGGATCAGCGACCCCGTGTGCTGTGACCTGTTCATTATTTGGCGGTTATCATACGAA	660
Db	601	CGGATCAGCGACCCCGTGTGCTGTGACCTGTTCATTATTTGGCGGTTATCATACGAA	660
OY	661	CACCACTGCACCCGACGGTGCTTTGTGTGGCGCTGCCAGACCCGCAACCAAGGGGAGAC	720
Db	661	CACCACTGCACCCGACGGTGCTTTGTGTGGCGCTGCCAGACCCGCAACCAAGGGGAGAC	720
OY	721	ACCGCATGA 729	
Db	721	ACCGCATGA 729	

RESULT 3
AX701576

DEFINITION	Sequence 180 from Patent WO20090905.
ACCESSION	AX701576
VERSION	AX701576.1 GI:29537147

SOURCE ORGANISM	Source
<i>Paracoccus carotinifaciens</i>	Paracoccus carotinifaciens
<i>Paracoccus carotinifaciens</i>	Paracoccus carotinifaciens
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;	

REFERENCE
AUTHORS
1
Berry, A., Bretzel, W., Huemelin, M., Lopez-Ulbarri, R., Mayer, A.F.
and Yeliseev, A.

JOURNAL
Patent: WO 0209095-A 180 12-DEC-2002;
Roche Vitamins AG (CH)
Location/Qualifiers

```

/organism="Paracoccus carotinifaciens"
/mol_type="unassigned DNA"
/db_xref="taxon:65151"

```

```

CDS
1..729
/note="unamed protein product"
/codon_start=1

```

```
/c/dms/_ladle=11  
/protein_id="CAD87641.1"  
/db_xref="GI:29537148"  
//seq_1+aa=IMCQVAPDPTATCTTSCGTTADPILHMLWEIDAAAD
```

LAVANFAGLIMWSVGLFI IARIAMHSSV FPGSRKANAMGQLVLM LINGF SIKRMIT
HMAHHRHAGTDDDPDEFDHGGPYRWARF IGTFYGMREGILLPIVITVVALMIGDRMM

ORIGIN

VFWPPLSLIASIOLFVFGIMLPHRPHGDAFPDRHNRSSRISDPVSLITCFHFGSYH
HEHHLPYPMWRILPSTRKGDPA"

Query Match 100.0%; Score 729; DB 6; Length 729;
Best Local Similarity 100.0%; Pred. No. 3.3e-110;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGGGCAATGCGCTTGGCCCAAGGAGATCTGACCCGACCGCTTGTATCGCTCGGGC 60
DB 1 ATGAGGGCAATGCGCTTGGCCCAAGGAGATCTGACCCGACCGCTTGTATCGCTCGGGC 60
QY 61 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 120
DB 61 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 120
QY 121 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 180
DB 121 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 180
QY 181 TTCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 240
DB 181 TTCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 240
QY 241 GCGGCGATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 300
DB 241 GCGGCGATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 300
QY 301 GTCAGACATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 360
DB 301 GTCAGACATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 360
QY 361 GCGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 420
DB 361 GCGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 420
QY 421 CTGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 480
DB 421 CTGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 480
QY 481 GTGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 540
DB 481 GTGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 540
QY 541 TGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 600
DB 541 TGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 600
QY 601 CGGATCAGGAGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 660
DB 601 CGGATCAGGAGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 660
QY 661 CACCACTGACCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 720
DB 661 CACCACTGACCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 720
QY 721 ACCGATGA 729
DB 721 ACCGATGA 729
```

RESULT 4
LOCUS AR169854 1261 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 30 from patent US 6291204.
ACCESSION AR169854
VERSION AR169854.1 GI:17907785
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1261)
AUTHORS Pasamontes, L. and Tsygankov, Y.

TITLE Fermentative carotenoid production
JOURNAL Patent: US 6291204-A-20-18-SEP-2001;
FEATURES location/Qualifiers
source 1.1261
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 729; DB 6; Length 1261;
Best Local Similarity 100.0%; Pred. No. 2.9e-110;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGGGCAATGCGCTTGGCCCAAGGAGATCTGACCCGACCGCTTGTATCGCTCGGGC 60
DB 40 ATGAGGGCAATGCGCTTGGCCCAAGGAGATCTGACCCGACCGCTTGTATCGCTCGGGC 99
QY 61 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 120
DB 100 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 159
QY 121 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 180
DB 160 GGCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 219
QY 181 TTCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 240
DB 220 TTCATATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 279
QY 241 GCGGCGATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 300
DB 280 GCGGCGATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 339
QY 301 GTCAGACATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 360
DB 340 GTCAGACATGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 399
QY 361 GCGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 420
DB 400 GCGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 459
QY 421 CTGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 480
DB 460 CTGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 519
QY 481 GTGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 540
DB 520 GTGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 579
QY 541 TGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 600
DB 580 TGGCTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 639
QY 601 CGGATCAGGAGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 660
DB 640 CGGATCAGGAGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 699
QY 661 CACCACTGACCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 720
DB 700 CACCACTGACCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 759
QY 721 ACCGATGA 729
DB 760 ACCGATGA 768
```

RESULT 5
LOCUS AR452028 1261 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 30 from patent US 6677134.
ACCESSION AR452028
VERSION AR452028.1 GI:42683355
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1261)
AUTHORS Paramontes, L. and Tsygankov, Y.
TITLE Fermentative carotenoid production
JOURNAL Patent: US 6677134-A 30 13-JAN-2004;
FEATURES Location/Qualifiers
source 1..1261
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 729; DB 6; Length 1261;
Best Local Similarity 100.0%; Pred. No. 2.9e-110;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGCGCATGCTCCCTGGCCCAAGGAGATCTGACCCGCAACGATTGATCGCTCCGGGC 60
DB 40 ATGAGCGGACATGCTCCCTGGCCCAAGGAGATCTGACCCGCAACGATTGATCGCTCCGGGC 99
QY 61 GGCATCATCGCGCGCTGGCGCTGGCCCTGATGTCATGTCATGCTGCTGCTGATGACGCGCG 120
DB 100 GGCATCATCGCGCGCTGGCGCTGGCCCTGATGTCATGTCATGCTGCTGCTGATGACGCGCG 159
QY 121 GCGCATCCCATCTCTGGCGCTGGCGAATTTCTTGGGGCTGACCTGCTGCTGCTGCTG 180
DB 160 GCGCATCCCATCTCTGGCGCTGGCGAATTTCTTGGGGCTGACCTGCTGCTGCTGCTG 219
QY 181 TTGATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 220 TTGATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 279
QY 241 GCGGCGATGGGCGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 280 GCGGCGATGGGCGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 339
QY 301 GTCAAGACATGAGCCCATCATGCGCATGCGGAAACCGACGACGCCAGATTTCAGCAT 360
DB 340 GTCAAGACATGAGCCCATCATGCGCATGCGGAAACCGACGACGCCAGATTTCAGCAT 399
QY 361 GCGGCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 400 GCGGCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459
QY 421 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 460 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 481 GTGCTCTTCTGCGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 520 GTGCTCTTCTGCGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
QY 541 TGGCTGCGGACCG 600
DB 580 TGGCTGCGGACCG 639
QY 601 GCGATCAGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 640 GCGATCAGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699
QY 661 CACCACTTGACCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 700 CACCACTTGACCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
QY 721 ACCGATGA 729
DB 760 ACCGATGA 768

RESULT 6
CQ788339 CQ788339 729 bp DNA linear PAT 24-MAR-2004
LOCUS CQ788339
DEFINITION Sequence 9 from Patent WO2004018695.

ACCESSION CQ788339
VERSION CQ788339.1 GI:45723106
KEYWORDS
SOURCE
ORGANISM Paracoccus marcusii
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;
Rhodobacteriaceae; Paracoccus.

ORIGIN

Query Match 95.8%; Score 698.6; DB 6; Length 729;
Best Local Similarity 97.4%; Pred. No. 3.3e-105;
Matches 710; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ATGAGCGACATGCTCCCTGGCCCAAGGAGATCTGACCCGCAACGATTGATGCTCCGGGC 60
DB 1 ATGAGCGGACATGCTCCCTGGCCCAAGGAGATCTGACCCGCAACGCTGATGCTCCGGGC 60
QY 61 GGCATCATGCGCGGCTGGCGCTGGCCCTGATGTCATGTCATGCTGCTGCTGATGACGCGCG 120
DB 61 GGCATCATGCGCGGCTGGCGCTGGCCCTGATGTCATGTCATGCTGCTGCTGATGACGCGCG 120
QY 121 GCGCATCCCATCTCTGGCGCTGGCGAATTTCTTGGGGCTGACCTGCTGCTGCTGCTG 180
DB 121 GCGCATCCCATCTCTGGCGCTGGCGAATTTCTTGGGGCTGACCTGCTGCTGCTGCTG 180
QY 181 TTGATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 TTGATCATGCGCGATGACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 GCGGCGATGGGCGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 241 GCGGCGATGGGCGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 241 GCGGCGATGGGCGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 301 GTCAAGACATGAGCCCATCATGCGCATGCGGAAACCGACGACGCCAGATTTCAGCAT 360
QY 301 GTCAAGACATGAGCCCATCATGCGCATGCGGAAACCGACGACGCCAGATTTCAGCAT 360
DB 301 GTCAAGACATGAGCCCATCATGCGCATGCGGAAACCGACGACGCCAGATTTCAGCAT 360
QY 361 GCGGCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GCGGCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GTGCTCTTCTGCGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GTGCTCTTCTGCGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 TGGCTGCGGACCG 600

Db	541	TGGCTGCGCGCAACCGCCCGCGGACAGAGAGCGTTCGCCGAGACCGCATTAAGGCGGGTCTGCG	600
Qy	601	CGGATCAGCGGACCCCGTGTCCGTGCTGACCTGCTTTCACTTTGGCGGTTATCATCAGCA	660
Db	601	CGGATCAGCGGACCCCGTGTCCGTGCTGACCTGCTTTCACTTTGGCGGTTATCATCAGCA	660
Qy	661	CACCAACCTGCACCCGAGCGGTGCTTGTGGGCGGCGCTGCCACACCCGACCAAGGAGGAGC	720
Db	661	CACCAACCTGCACCCGAGCGGTGCTTGTGGGCGGCGCTGCCACACCCGACCAAGGAGGAGC	720
Qy	721	ACCGCATGA 729	
Db	721	ACCGCATGA 729	
RESULT 7			
LOCUS	CQ793306	729 bp DNA	Linear PAT 29-MAR-2004
DEFINITION	Sequence 9 from Patent WO2004017749.		
ACCESSION	CQ793306		
VERSION	CQ793306.1	GI:45824403	
KEYWORDS			
SOURCE	Paracoccus marcusii		
ORGANISM	Paracoccus marcusii		
REFERENCE	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Paracoccus.		
AUTHORS	1. Pflichmann, R., Sauer, M., Schopfer, C.R., Klebsattel, M., Pfeiffer, A.M., Luck, T. and Voeste, D.		
TITLE	Use of asetraxanthin-containing plants or parts of plants of the genus tagetes as animal feed		
JOURNAL	Patent: WO 2004017749-A 9 04-MAR-2004; Sungene GmbH & Co. KGAA (DE); BASF Aktiengesellschaft (DE); BASF Plant Science GmbH (DE)		
FEATURES	Location/Qualifiers		
SOURCE	1..729		
CDS	/organism="Paracoccus marcusii"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:59779"		
	1..729		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=1		
	/protein_id="CAG24515.1"		
	/db_xref="GI:45824404"		
	/translation="MSAHLPEKADLTATSLIVSGIITAAWLALHVALMFLDAAPHITLAVNPLGLTWLSTGLFIADHDAHMGSLVPRPRANAMGOLVILVLAGSWMRIYKRMHAIHRNAGCTDDDDPDHGGGVRVYARPICTYRPMRELLPVIYVYAIILGRMRYVVFNPUSITIASIQLFVFGTLPRIHPGHDAFPDRHNASSRISDPVSILITCFHFGYTH		
	HEHHLHPVPMWRLPSTRTKDTA"		
ORIGIN			
Query Match	95.8%; Score 698.6; DB 6; Length 729;		
Best Local Similarity	97.4%; Pred. No. 3.3e-105;		
Matches 710; Conservative	0; Mismatches 19; Indels 0; Gaps 0;		
Qy	1	ATGAGCGGACATGCCCCGCGGAGGAGGAGATGACCGGACACGATTGTGCTCGGGC	60
Db	1	ATGAGCGGACATGCCCCGCGGAGGAGGAGATGACCGGACACGATTGTGCTCGGGC	60
Qy	61	GCGATCATGCGCGGTGGCTGAGCTTCCTGCGGGCTGACCTGGCTGCGGTGCTG	120
Db	61	GCGATCATGCGCGGTGGCTGAGCTTCCTGCGGGCTGACCTGGCTGCGGTGCTG	120
Qy	121	GCGATCATGCGCGGTGGCTGAGCTTCCTGCGGGCTGACCTGGCTGCGGTGCTG	180
Db	121	GCGATCATGCGCGGTGGCTGAGCTTCCTGCGGGCTGACCTGGCTGCGGTGCTG	180
Qy	181	TTTCATCATGCGGACGAGATGATGAGTCTGCTGCGGAGCGCGCGGCGCAAT	240
Db	181	TTTCATCATGCGGACGAGATGATGAGTCTGCTGCGGAGCGCGCGGCGCAAT	240
Qy	241	GCGGCGATGGCGACGCTTGTCTGCTGCTGTATGCGGATTTTCTTGCGCGCAATGATC	300

Query Match	Score	698.6	DB 6	Length	729
Db	241	CGCGCGATGGGCGAGCTTGTCTGTGGCTGTATGCCGATTTTGTGGCCAGATGATC	300		
Qy	301	GTCAAGCAATGCGCCATCATCGCCATGCGGAAACGACGACGACCAAGATTTGACCAT	360		
Db	301	GTCAAGCAATGCGCCATCATCGCCATGCGGAAACGACGACGACCAAGATTTGACCAT	360		
Qy	361	GGCGGCGCGGTCCGTGTGTATGCCCGCTTATATGCGGACCTATTGGCTGGCGGAGGG	420		
Db	361	GGCGGCGCGGTCCGTGTGTATGCCCGCTTATATGCGGACCTATTGGCTGGCGGAGGG	420		
Qy	421	CTGTGTCTGCGCCGTCATCTGTATGCGCTATATGTTGGGGGATTCGTGATGTAC	480		
Db	421	CTGTGTCTGCGCCGTCATCTGTATGCGCTATATGTTGGGGGATTCGTGATGTAC	480		
Qy	481	GTGTCTTCTTGCGCGGTGCGGTGCGATCTGGCGTCATGACGCTGTCGTGTGCGATC	540		
Db	481	GTGTCTTCTTGCGCGGTGCGGTGCGATCTGGCGTCATGACGCTGTCGTGTGCGACT	540		
Qy	541	TGCGTCCGCGACCGCGCCCGGCGACGACGCGTTCGCCGACATGCGCGGTCTGTG	600		
Db	541	TGCGTCCGCGACCGCGCCCGGCGACGACGCGTTCGCCGACATGCGCGGTCTGTG	600		
Qy	601	CGATTCAGCGACCGCGGTGCGCTGTACCTGCTTCACTTTGGCGGTTATCATCAGAA	660		
Db	601	CGATTCAGCGACCGCGGTGCGCTGTACCTGCTTCACTTTGGCGGTTATCATCAGAA	660		
Qy	661	CACCACTGCACCGCGAGTGCCTTGTGTGGCGCTGCGCCACGACCGCACCAAGGGGAC	720		
Db	661	CACCACTGCACCGCGAGTGCCTTGTGTGGCGCTGCGCCACGACCGCACCAAGGGGAC	720		
Qy	721	ACCGCATGA 729			
Db	721	ACCGCATGA 729			
RESULT 8					
LOCUS	CQ793487				
DEFINITION	Sequence 9 from Patent WO2004018693.				
ACCESSION	CQ793487				
VERSION	CQ793487.1				
KEYWORDS	GI:45824578				
SOURCE					
ORGANISM	Paracoccus marcusii				
REFERENCE	1				
AUTHORS	Schopfer, C.R., Flachmann, R., Hebers, K., Kunze, I., Sauer, M. and Klebsattel, M.				
TITLE	Method for the production of ketocarotins in flower petals on plants				
JOURNAL	Patent: WO 2004018693-A 9 04-MAR-2004;				
FEATURES					
source	Location/Qualifiers				
	1..729				
	/organism="Paracoccus marcusii"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:59779"				
	1..729				
	/note="unnamed protein product"				
	/codon_start=1				
	/transl_table=1				
	/protein_id="CAG24548.1"				
	/db_xref="GI:45824579"				
	/translation="MSAHLPKADITATSIIVSGGIIAHLALHVAHFLDAAHPI				
	LAHAFIAGLTWLSVGLFTLHDMHGSVYGRPRANAMQVLVLAHGSRRKMTYK				
	LMHHHRHAGTDDDDPDHGGPVRWYAFPGTGYGMRGGLLPVITVYALILDGRMTY				
	VVWVLPSTIASIQLFVFGTMLPHRPGHDAPFDRHNRSSRIDPVSLLTCTFHGGYH				
	HEHHLHPTVPMWRLPSTRKGDTA"				
ORIGIN					


```

/protein_id="CAB56062.1"
/db_xref="GI:5912295"
/db_xref="GOA:O9RLH4"
/db_xref="UniProt/TREMBL:O9RLH4"
/translation="MNAHSPAKTAVIVAGFEGCALAIRLOSAGIATTLVBARXPG
GRAVYHDDQGHVFDAGPTVITDPDALKEMLALTGDMADVTLMVSPRYRLMWPBGK
VADYVEADQLESQIAQNPDDLEGRRADYAEVYQGYVLAQVPLKIGOMLKA
AALMLLEAVNSVHAKVATEIKDPYLRQAFSYHTLVGNGPSTSIYALIHLEBRG
GWPFAKGTNOLVAGMVALFERPGCOMINAVARIETSGATGTGTLADGSLRDM
VASNGDVHNTNRDLGHTARGOSRAKSLDRKMSLSPLVHGLGRAPADIAHTILF
GRYKELVNEIFRGPCLADFSLYKHSPTCTPDMAPEMSTHYVLAIPHLISRAID
WAVEGPRYADRIALFLERLIPTLRANLIRTRIPTPADASELNAHSGAFSEVPIIT
OSAFEPHNRDITIRNFYLVGATGHPGAGIPGVGSAAKATAQVMSDLASA"
4278..5192
/gene="crtb"
4278..5192
/gene="crtb"
/function="astaxanthin biosynthesis"
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="CAB56063.1"
/db_xref="GI:5912296"
/db_xref="GOA:O9RLH3"
/db_xref="UniProt/TREMBL:O9RLH3"
/translation="MSDVLVSTSEATIGOSOPATPAKILMPGIRDDVTMLYAMCRHA
DIVIGQALGSRPEANVDPOARLDGLRDTALAGDGVTPPRALVABRRHBPQ
AMPMDLIEGFALDVEARBYRTLDVLEYSYHVAIGLVMMKAVGRDHPVLDACDL
GLAFDITRIARVIDYARIGRCYLPEDMDLHAGARVDGVPSELTIVLIRLDAEP
YPSARVGLADLPPRCAMSI PAFRYRVLIGLPIRKGSEAVRORISTSKAKIGLIG
IGMPVADHACGSGSVRODLWTRPHHA"
complement(5189..6064)
/gene="crtb"
complement(5189..6064)
/gene="crtb"
/codon_start=1
/transl_table=11
/product="geranylgeranyl pyrophosphate synthase"
/protein_id="CAB56064.1"
/db_xref="GI:5912297"
/db_xref="GOA:O9RLH2"
/db_xref="UniProt/TREMBL:O9RLH2"
/translation="MRDYNPIHATLORIEIAGQSEAVLQPIGAAMALSGKRF
KEMMLIABASGVCDTIVDAACAVEMHAASLIFDDPCMDDAIRGQATRYAH
GSSRAVGLGILITEMMALAARGASGVRAQVLTLSRSJGPOGLCAQDULDAHA
KNGAGVEQDQKDTGVLFYAGLEMLAVIKPEAEQTMIDGRQGRVFOGYDILIE
IVGDQALGKDTGRDAAAPGPRRGLIADVLDQKVSROYASRAQLEGMVSKRLQAPE
IALLEKRVLPYARA"

```

ORIGIN

```

Query Match      95.8%; Score 698.6; DB 1; Length 6335;
Best Local Similarity 97.4%; Pred. No. 2e-105;
Matches 710; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 AAGAGGCAATGCCCTGCCCAAGAGATCGACCGCAACGATTGATCGCTCGGGC 60
DB 409 AAGAGGCAATGCCCTGCCCAAGAGATCGACCGCAACGATTGATCGCTCGGGC 468
QY 61 GGCATCATGCCGCGGTGCTGGCCCTGCATGTCATGCCCTGCTGTCGACGCGGCG 120
DB 469 GGCATCATGCCGCGGTGCTGGCCCTGCATGTCATGCCCTGCTGTCGACGCGGCG 528
QY 121 GCGCATCCCATCTGCGGCTCGGAATTTCTGCGGCTGACCTGCGCTGCGGCTG 180
DB 529 GCGCATCCCATCTGCGGCTCGGAATTTCTGCGGCTGACCTGCGCTGCGGCTG 588
QY 181 TTCATCATGCCGATGACGCGATGATGATGCTGCTGCTGCGGCGCGCGCGCAT 240
DB 589 TTCATCATGCCGATGACGCGATGATGATGCTGCTGCTGCGGCGCGCGCGCAT 648
QY 241 GCGGCGATGGCGAGCTTCTGCTGCTGATGCGCGATTTTCTGCGCGCAAGATGATC 300
DB 241 GCGGCGATGGCGAGCTTCTGCTGCTGATGCGCGATTTTCTGCGCGCAAGATGATC 300

```

```

DB 649 GCGGCGATGGCGAGCTTCTGCTGCTGATGCGCGATTTTCTGCGCGCAAGATGATC 708
QY 301 GTCAGGCAATGGCCCATCATGCGCATGCGGAGACCGAGACCGCAATTTGACCAT 360
DB 709 GTCAGGCAATGGCCCATCATGCGCATGCGGAGACCGAGACCGCAATTTGACCAT 768
QY 361 GCGGCGCGCGTCCGCTGTGACCCCGGCTTCATCGGCACTATTTGCGTGGCGAGGGG 420
DB 769 GCGGCGCGCGTCCGCTGTGACCCCGGCTTCATCGGCACTATTTGCGTGGCGAGGGG 828
QY 421 CTGCTGCTGCCGCTCATGCTGACCGCTCTATGCGCTGATTTGGGGGATGCTGATGAC 480
DB 829 CTGCTGCTGCCGCTCATGCTGACCGCTCTATGCGCTGATTTGGGGGATGCTGATGAC 888
QY 481 GTGCTCTTCTGCGCGCTGCGCTGATGCTGCGGAGTGCATGCTGCTGCTGCGGATC 540
DB 889 GTGCTCTTCTGCGCGCTGCGCTGATGCTGCGGAGTGCATGCTGCTGCTGCGGATC 948
QY 541 TGGCTGCCGACACCGGCGCGGCGGACGAGCGGTTCCGGAACCGGCAATGCGGCTGTCG 600
DB 949 TGGCTGCCGACACCGGCGCGGCGGACGAGCGGTTCCGGAACCGGCAATGCGGCTGTCG 1008
QY 601 CGGATGACGACCGCGGTGCTGCTGCTGACCTGCTTCACTTTGGCGGTTATCATGACGA 660
DB 1009 CGGATGACGACCGCGGTGCTGCTGCTGACCTGCTTCACTTTGGCGGTTATCATGACGA 1068
QY 661 CACCACTGACACCGGACCGGCTGCTGCGGCGGCTGCGGACCGGCAATGCGGAGGGGAC 720
DB 1069 CACCACTGACACCGGACCGGCTGCTGCGGCGGCTGCGGACCGGCAATGCGGAGGGGAC 1128
QY 721 ACCGATGAC 729
DB 1129 ACCGATGAC 1137

```

```

RESULT 10
LOCUS CQ788335 729 bp DNA linear PAT 24-MAR-2004
DEFINITION Sequence 5 from Patent WO2004018655.
ACCESSION CQ788335
VERSION CQ788335.1 GI:45723102
KEYWORDS
SOURCE
ORGANISM Agrobacterium aurantiacum
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE
1 Schopfer,C.R., Flachmann,R., Herbers,K., Kunze,I., Sauer,M. and
Klebsattel,M.
TITLE Method for producing ketocarotenoids in plant fruit
JOURNAL Patent: WO 2004018655-A 5 04-MAR-2004;
Sungene GmbH & Co. KGaA (DE)
FEATURES
source
1..729
location/Qualifiers
/organism="Agrobacterium aurantiacum"
/mol_type="unasigned DNA"
/db_xref="taxon:44155"
1..729
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAG17901.1"
/translation="MSAIALPRAADITATSLIVSGGIIAAMLAALHVALWFLDAAPPI
LAIANFLGLTWLSVGLFIADHAMGSVPGSRPRANAMGDLVWLVAGFSRKMIVK
HMAHRRHAGTDDDPDGHGQVRAWYARIGYFGREGLLPLVITVVALILIGDRMY
VVFMPLESLIASIOLFVFGTWLPHRPGHADPDRHNAASRISDPVSLTCTPHFGYH
HEHHLHPVPMRLPSTRTKDTA"

```

ORIGIN

```

Query Match      95.6%; Score 697; DB 6; Length 729;
Best Local Similarity 97.3%; Pred. No. 6.1e-105;
Matches 709; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

OY		1	ATGAGCGCAATGCCCTCCGAAGAACAATGTGACCGGCACAGATTATCATGGTCTCGGGC	60
Db		1	ATGAGCGCAATGCCCTCCGAAGAACAATGTGACCGGCACAGATTATCATGGTCTCGGGC	60
OY		61	GGCATCATCGCCGCGTGAGCTGGCCCTTCGATGTGCATGCGCTGTGGTTTCTGGACGCGCG	120
Db		61	GGCATCATCGCCGCGTTGGCTGGCCCTTCGATGTGCATGCGCGCTGTGGTTTCTGGACGCGCG	120
OY		121	GCGCATGCCATTCCTGAGCGGTGGCGAAATTTCCTGGGGGCTGACCTGGCTGTGGGTGGTCTG	180
Db		121	GCGCATGCCATTCCTGAGCGGTGGCGAAATTTCCTGGGGGCTGACCTGGCTGTGGGTGGATTTG	180
OY		181	TTCATCATCGCGCATGACGCGATGATGGGTGCGGTGCGCGGGCGCCCGCGCCCAAT	240
Db		181	TTCATCATCGCGCATGACGCGATGACGCGGTGCGGTGCGCGGGCGTCCGCGCCCAAT	240
OY		241	GCGGCGATGGGCGACGTTGTCTGTGGCTGTATGCGCGAATTTTCTGCGCGCAAGATGATC	300
Db		241	GCGGCGATGGGCGACGTTGTCTGTGGCTGTATGCGCGAATTTTCTGCGCGCAAGATGATC	300
OY		301	GTCAGACATATGCCCCATCATTCGCCATGCCCAGAACCGACGACGACCAGATTTTGCACCAT	360
Db		301	GTCAGACATATGCCCCATCATTCGCCATGCCCAGAACCGACGACGACCAGATTTTGCACCAT	360
OY		361	GAGCGCCCGGTGCGCGTGGTAGCGCCCGCTTCATCGSACCTTAATTTTGGCTGGGCGCGAGGG	420
Db		361	GAGCGCCCGGTGCGCGTGGTAGCGCCCGCTTCATCGSACCTTAATTTTGGCTGGGCGCGAGGG	420
OY		421	CTGCTGTGCCCCGTATGATGTGACGGTGTATGCGCTGATGTTGGGGGATCGCTGATGTAC	480
Db		421	CTGCTGTGCCCCGTATGATGTGACGGTGTATGCGCTGATGTTGGGGGATCGCTGATGTAC	480
OY		481	GTTGGCTTCTGGCGGTTCGCCGTGCATCTCGGCGTGCATCGACGTGGTGGTGGTGGCGATC	540
Db		481	GTTGGCTTCTGGCGGTTCGCCGTGCATCTCGGCGTGCATCGACGTGGTGGTGGTGGCGATC	540
OY		541	TGGCTGCGCGCACCGCCCGCGGCACACGACGCGTCCCGGACCGGCACATGCGCGGTGCTCG	600
Db		541	TGGCTGCGCGCACCGCCCGCGGCACACGACGCGTCCCGGACCGGCACATGCGCGGTGCTCG	600
OY		601	CGGATCAGCGACCCCGGTGTGCTGTGACCTGTCTTTCATCTTTGGCGGTTATCATCAGAA	660
Db		601	CGGATCAGCGACCCCGGTGTGCTGTGACCTGTCTTTCATCTTTGGCGGTTATCATCAGAA	660
OY		661	CACCACTCGACCCGACGCGTGTGGTGGCGGCTGCCACGACCCGACACCAAAGGGGAC	720
Db		661	CACCACTCGACCCGACGCGTGTGGTGGCGGCTGCCACGACCCGACACCAAAGGGGAC	720
OY		721	ACCGCATGA	729
Db		721	ACCGCATGA	729
RESULT	11			
LOCUS	CQ788369			
DEFINITION	Sequence 39 from Patent WO2004018695.	729 bp	DNA	linear PAT 24-MAR-2004
ACCESSION	CQ788369			
VERSION	CQ788369.1	GI:45723136		
KEYWORDS				
SOURCE				
ORGANISM	Paracoccus sp. MBIC1143 Paracoccus sp. MBIC1143 Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales; Rhodobacteraceae; Paracoccus.			
REFERENCE				
AUTHORS	Schopfer,C.R., Flachmann,R., Herbers,K., Kunze,I., Sauer,M. and Klabsattel,M.			
JOURNAL	Method for producing ketocarotenoids in plant fruit Patent: WO 2004018695-A 39 04-MAR-2004;			
FEATURES	Sungene GmbH & Co. KGAA (DE) Location/Qualifiers			
SOURCE		1..729		

CD5

1. 729
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA91.7910.1"
/db_xref="GI:45723137"
/translation="MSAHALPKADLTATSLIVSGGIIIAAWLALHVAHLMFLDAAHPI
LAINSLGLTWLSVGLFTIAHDAHSGSVVGRPRANAAMQULYLMYAGSRKMTYK
HMAHHRAGTDDDDFDHGGVPRVYTAIFIGYFEDRWKGLLPIVLYTALILGRMT
VVFPEPLSIILASIQLFVFGTWLPKRPDHAFFPRHNARSRIISDPVSLTICFHHGVYH
HEHHLHPTVPMWRLLPSTRTKGDTA"

ORIGIN

Query Match	95.6%	Score 697;	DB 6;	Length 729;
Best Local Similarity	97.3%	Pred. No. 16-105;		
Matches 709;	Conservative 0;	Mismatches 20;	Indels 0;	Gaps 0;
QY	1	ATGAGCCACATGCGCTGCGCCCAAGGAGATGTGACCCGCCACAGTTTGATGTCGCGGC	60	
Db	1	ATGAGCCACATGCGCTGCGCCCAAGGAGATGTGACCCGCCACAGTTTGATGTCGCGGC	60	
QY	61	GGCATCATGCGCGGTGAGCTGCGCCCTGCATGTGCATGGCTGTGTGTTCTGGAAGCGGCG	120	
Db	61	GGCATCATGCGCGGTGAGCTGCGCCCTGCATGTGCATGGCTGTGTGTTCTGGAAGCGGCG	120	
QY	121	GGCATCCCATCTGCGGAGATGCAAATTTCTGAGGCTGACCTGGCTGTGCGGTGCGATTG	180	
Db	121	GGCATCCCATCTGCGGAGATGCAAATTTCTGAGGCTGACCTGGCTGTGCGGTGCGATTG	180	
QY	181	TTTCATCATGCGGATGACGCGATGCATGGGTGCGTCCGCGGCGCCCGGCCAAT	240	
Db	181	TTTCATCATGCGGATGACGCGATGCATGGGTGCGTCCGCGGCGCCCGGCCAAT	240	
QY	241	GGCGCGATGCGGCGCGCTTGTCTGTGTGTGTATGACCGAATTTTCTGCGCGCAATATC	300	
Db	241	GGCGCGATGCGGCGCGCTTGTCTGTGTGTGTATGACCGAATTTTCTGCGCGCAATATC	300	
QY	301	GTCAAGACATGAGCCCATCATGCGCATGCGGAGACGACGACGACCGACAGTTTGCACAT	360	
Db	301	GTCAAGACATGAGCCCATCATGCGCATGCGGAGACGACGACGACCGACAGTTTGCACAT	360	
QY	361	GGCGCGCGGTGCGCGTGTGACGCGCGCTTCACTGCGCACTATTTGCGTGTGCGCGAGGCG	420	
Db	361	GGCGCGCGGTGCGCGTGTGACGCGCGCTTCACTGCGCACTATTTGCGTGTGCGCGAGGCG	420	
QY	421	CTGCTGTGCGCGCGCATGATGATGAGCGGTCTATGCGCTGATGTTGGGAGATGCGTGTATGAC	480	
Db	421	CTGCTGTGCGCGCGCATGATGATGAGCGGTCTATGCGCTGATGTTGGGAGATGCGTGTATGAC	480	
QY	481	GTGTGCTTCTTGCGCGGTGCGGTGATCTGAGCGTGCATTCAGACTGTTCTGTGTCGCAATC	540	
Db	481	GTGTGCTTCTTGCGCGGTGCGGTGATCTGAGCGTGCATTCAGACTGTTCTGTGTCGCAATC	540	
QY	541	TGGGTGCGCGACCGCGCGCGCGCAAGAGCGGTTCCCGAGACCGCAATGCGCGGTGCGCG	600	
Db	541	TGGGTGCGCGACCGCGCGCGCGCAAGAGCGGTTCCCGAGACCGCAATGCGCGGTGCGCG	600	
QY	601	CGATTCAGCGACCCCGTGTGCTGTGCTGACCTTTCATCTTGGCGGTTATCATCAGAA	660	
Db	601	CGATTCAGCGACCCCGTGTGCTGTGCTGACCTTTCATCTTGGCGGTTATCATCAGAA	660	
QY	661	CACCACTGACCCGAGCGGTGCTTGGTGGGCGCTGCCACGACCCGACCAAGGGGAC	720	
Db	661	CACCACTGACCCGAGCGGTGCTTGGTGGGCGCTGCCACGACCCGACCAAGGGGAC	720	
QY	721	ACCGCATGA 729		
Db	721	ACCGCATGA 729		

RESULT 12
 CQ793302 729 bp DNA linear PAT 29-MAR-2004
 LOCUS Sequence 5 from Patent WO2004017749.
 DEFINITION CQ793302
 ACCESSION CQ793302.1 GI:45824399
 VERSION
 KEYWORDS
 SOURCE Agrobacterium aurantiacum
 ORGANISM Agrobacterium aurantiacum
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobaceae; Rhizobium/Agrobacterium group; Agrobacterium.

REFERENCE
 1 Pfachmann, R., Sauer, M., Schopfer, C.R., Klebsattel, M., Pfeiffer, A.M., Luck, T. and Voeste, D.
 Use of astaxanthin-containing plants or parts of plants of the genus tagetes as animal feed
 Patent: WO 2004017749-A 5 04-MAR-2004;
 Sungene GmbH & Co. KGaA (DE); BASF Aktiengesellschaft (DE); BASF Plant Science GmbH (DE)

JOURNAL
 Location/Qualifiers
 1..729
 /organism="Agrobacterium aurantiacum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:44155"
 1..729
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAG24513.1"
 /db_xref="GI:45824400"
 /translation="MSAHALPKADLTATSLIVSGIIAAMLAHLVHAFMDAAHPILAIANFLGLTWLSVGLFIADHMGSVVGRPRANAAGQVLMLYAFSWRMKIVKHNAAHHAAGTDDPDPDHGSPVRYARFICTYGRMGEGILLPIVIVYVALLIGDRMYVYVWPLPSILASIQLEFVFGTWLPHRPHGDAFDRHNAARSRLSDPVSLLTCFHFSGYHHEHHHLPTVPMWRLPSTRTKGDTA"

ORIGIN
 Query Match 95.6%; Score 697; DB 6; Length 729;
 Best Local Similarity 97.3%; Pred. No. 6.1e-105;
 Matches 709; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

DB 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 QY 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 DB 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 QY 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180
 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180
 DB 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180
 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180
 QY 181 TTCATATGCGCGATGACGCGATGACGCGATGACGCGATGACGCGATGACGCGAT 240
 181 TTCATATGCGCGATGACGCGATGACGCGATGACGCGATGACGCGATGACGCGAT 240
 DB 181 TTCATATGCGCGATGACGCGATGACGCGATGACGCGATGACGCGATGACGCGAT 240
 181 TTCATATGCGCGATGACGCGATGACGCGATGACGCGATGACGCGATGACGCGAT 240
 QY 241 GCGCGATGCGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 300
 241 GCGCGATGCGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 300
 DB 241 GCGCGATGCGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 300
 241 GCGCGATGCGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 300
 QY 301 GTCAAGCAATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 360
 301 GTCAAGCAATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 360
 DB 301 GTCAAGCAATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 360
 301 GTCAAGCAATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCGCGATGCG 360
 QY 361 GCGCGCGCGCGTCCGCTGCTAGCGCGCGCTTATCGGCACTATTTCCGCTGGCGG 420
 361 GCGCGCGCGCGTCCGCTGCTAGCGCGCGCTTATCGGCACTATTTCCGCTGGCGG 420
 DB 361 GCGCGCGCGCGTCCGCTGCTAGCGCGCGCTTATCGGCACTATTTCCGCTGGCGG 420
 361 GCGCGCGCGCGTCCGCTGCTAGCGCGCGCTTATCGGCACTATTTCCGCTGGCGG 420
 QY 421 CTGCTGCTGCGCTCATGTCAGCGTCTATGCGCTGATGTTGGGGGATCGCTGATGAT 480
 421 CTGCTGCTGCGCTCATGTCAGCGTCTATGCGCTGATGTTGGGGGATCGCTGATGAT 480
 DB 421 CTGCTGCTGCGCTCATGTCAGCGTCTATGCGCTGATGTTGGGGGATCGCTGATGAT 480
 421 CTGCTGCTGCGCTCATGTCAGCGTCTATGCGCTGATGTTGGGGGATCGCTGATGAT 480

QY 481 GTGCTCTTCTGCGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 540
 481 GTGCTCTTCTGCGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 540
 DB 481 GTGCTCTTCTGCGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 540
 481 GTGCTCTTCTGCGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCTTCTGCGCT 540
 QY 541 TGCTGCTGCG 600
 541 TGCTGCTGCG 600
 DB 541 TGCTGCTGCG 600
 541 TGCTGCTGCG 600
 QY 601 CGGATCAGGAGCG 660
 601 CGGATCAGGAGCG 660
 DB 601 CGGATCAGGAGCG 660
 601 CGGATCAGGAGCG 660
 QY 661 CACCACTGACCG 720
 661 CACCACTGACCG 720
 DB 661 CACCACTGACCG 720
 661 CACCACTGACCG 720
 QY 721 ACCGATGTA 729
 721 ACCGATGTA 729
 DB 721 ACCGATGTA 729
 721 ACCGATGTA 729

RESULT 13
 CQ793396 729 bp DNA linear PAT 29-MAR-2004
 LOCUS Sequence 99 from Patent WO2004017749.
 DEFINITION CQ793396
 ACCESSION CQ793396.1 GI:45824493
 VERSION
 KEYWORDS
 SOURCE Paracoccus sp. MB1C1143
 ORGANISM Paracoccus sp. MB1C1143
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;
 Rhodobacteraceae; Paracoccus.

REFERENCE
 1 Pfachmann, R., Sauer, M., Schopfer, C.R., Klebsattel, M., Pfeiffer, A.M., Luck, T. and Voeste, D.
 Use of astaxanthin-containing plants or parts of plants of the genus tagetes as animal feed
 Patent: WO 2004017749-A 99 04-MAR-2004;
 Sungene GmbH & Co. KGaA (DE); BASF Aktiengesellschaft (DE); BASF Plant Science GmbH (DE)

JOURNAL
 Location/Qualifiers
 1..729
 /organism="Paracoccus sp. MB1C1143"
 /mol_type="unassigned DNA"
 /db_xref="taxon:81397"
 1..729
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAG24528.1"
 /db_xref="GI:45824494"
 /translation="MSAHALPKADLTATSLIVSGIIAAMLAHLVHAFMDAAHPILAIANFLGLTWLSVGLFIADHMGSVVGRPRANAAGQVLMLYAFSWRMKIVKHNAAHHAAGTDDPDPDHGSPVRYARFICTYGRMGEGILLPIVIVYVALLIGDRMYVYVWPLPSILASIQLEFVFGTWLPHRPHGDAFDRHNAARSRLSDPVSLLTCFHFSGYHHEHHHLPTVPMWRLPSTRTKGDTA"

ORIGIN
 Query Match 95.6%; Score 697; DB 6; Length 729;
 Best Local Similarity 97.3%; Pred. No. 6.1e-105;
 Matches 709; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

DB 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 QY 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 DB 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 1 ATGAGGCAACATCCCTGCGCAAGCAGATCTGACCGCCACGATTGATCGTCTCGGGC 60
 QY 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 DB 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 61 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 120
 QY 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180
 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180
 DB 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180
 121 GGCATATGCGCGCGTGGCTGGCCCTGATGTCAGTGGCTGTTCTGACGCGGCG 180

REFERENCE

1 Schopfer, C.R., Flachmann, R., Herbers, K., Kunze, I., Sauer, M. and Klebatschel, M.

TITLE

Method for the production of ketocarotenoids in flower petals on plants

JOURNAL

Patent: WO 2004/018693-A 83 04-MAR-2004;

FEATURES

Sungene GmbH & Co. KGaA (DE)

Source

1. 729

/organism="Paracoccus sp. MBIC1143"

/mol_type="unassigned DNA"

/db_xref="taxon:81397"

1. 729

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAG24558.1"

/db_xref="GI:45824653"

/translation="MSAHLPKADLTATSLIVSGIIAAMLAHLVHMLPDAAHPI
LAIANFLGLTWLSVGLFLIADHMGSVVGRPRANAMGOLVLTLYAGSPWKMIVK
HMAHHRAGTDDPDPDHGGPRVWVARFVIGTQWREGILLPIVTVVALLIGDRMY
VVFWRPLSLIASIQLFVFGTWLPHRPHDAFPDRHNRSSRISDPVSLTCTHFGSIH
HEHHLHPVPMWRLPSTRTKGDYA"

ORIGIN

Query Match

Best Local Similarity 95.6%; Score 697; DB 6; Length 729;

Matches 709; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

QY 1 ATGAGGCGACATGCGCTGCGCCAGGAGATCTGACCGCCACAGTTTGATCGTCCGGC 60
   |||||
Db 1 ATGAGGCGACATGCGCTGCGCCAGGAGATCTGACCGCCACAGTTTGATCGTCCGGC 60

QY 61 GGCATCATGCGCGCGGTGCGTGGCTGCGATGTCATGCGCTGCTTCTGACGCGCG 120
   |||||
Db 61 GGCATCATGCGCGCGGTGCGTGGCTGCGATGTCATGCGCTGCTTCTGACGCGCG 120

QY 121 GGGCATCCATCTGCGGCGTGGAGATTCTGCGGCGTGAACCTGCGCTGCGTGGCTG 180
   |||||
Db 121 GGGCATCCATCTGCGGCGTGGAGATTCTGCGGCGTGAACCTGCGCTGCGTGGCTG 180

QY 181 TTCATCATGCGCGATGACGCGATGATGAGTCCGTGCGCGGCGCGCCGCGCCAAAT 240
   |||||
Db 181 TTCATCATGCGCGATGACGCGATGATGAGTCCGTGCGCGGCGCGCCGCGCCAAAT 240

QY 241 GCGGCGATGCGCGAGCTTGTCTGTGCGTGTATGCGGATTTTCTGCGGCGAAGATGATC 300
   |||||
Db 241 GCGGCGATGCGCGAGCTTGTCTGTGCGTGTATGCGGATTTTCTGCGGCGAAGATGATC 300

QY 301 GTCAAGCAGATGCGCGCATGATGCGCGATGCGCGCGGAAACGACGACCAATTTGACCAAT 360
   |||||
Db 301 GTCAAGCAGATGCGCGCATGATGCGCGATGCGCGCGGAAACGACGACCAATTTGACCAAT 360

QY 361 GGGGCGCGCGTCCGCTGTATGCGCGCGCTTTCATGCGACCTAATTTGCGTGGCGGAGGG 420
   |||||
Db 361 GGGGCGCGCGTCCGCTGTATGCGCGCGCTTTCATGCGACCTAATTTGCGTGGCGGAGGG 420

QY 421 CTGCTGCTGCGCGTATCTGTACGCTGTATGCGCTGATGTTGGGGGATGCTGATGTAC 480
   |||||
Db 421 CTGCTGCTGCGCGTATCTGTACGCTGTATGCGCTGATGTTGGGGGATGCTGATGTAC 480

QY 481 GTGGCTTTCGCGCGTGTGCGGATCTGCGCGTGTGATCCAGCTGTTGCGTTCGCGCATC 540
   |||||
Db 481 GTGGCTTTCGCGCGTGTGCGGATCTGCGCGTGTGATCCAGCTGTTGCGTTCGCGCATC 540

QY 541 TGGCTGCGCGACCGCGCGCGCGACGACGCGCTTCCGACCGCGACCAATGCGCGGTGCTG 600
   |||||
Db 541 TGGCTGCGCGACCGCGCGCGCGACGACGCGCTTCCGACCGCGACCAATGCGCGGTGCTG 600

QY 601 CGGATCAGCGACCGCGGTGCTGCTGATGCTGCTTTCATCTTGGCGGTTATCATCAGAA 660
   |||||
Db 601 CGGATCAGCGACCGCGGTGCTGCTGATGCTGCTTTCATCTTGGCGGTTATCATCAGAA 660

```

```

QY 661 CACCACTGACACCGGACGCGTGCCTTGGTGGCGCTGCGCCAGCAACCGGACCAAGGGGAGAC 720
   |||||
Db 661 CACCACTGACACCGGACGCGTGCCTTGGTGGCGCTGCGCCAGCAACCGGACCAAGGGGAGAC 720

QY 721 ACCGATGA 729
   |||||
Db 721 ACCGATGA 729

```

Search completed: June 21, 2005, 22:57:51
Job time : 3379.93 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 18:55:34 ; Search time 143.862 Seconds

(without alignments)
1359.734 Million cell updates/sec

Title: US-10-695-980-5

Perfect score: 1999

Sequence: 1 MSHDLLIAGAGLGGALIALA.....IPLSQAVRCLPRPRLQERA 382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	382	P94791	flavobacter
2	1456	72.8	386	CRTY_AGRAU	agrobacteri
3	1405	70.3	386	O9RLH5	paracoccus
4	810.5	40.5	382	CRTI_PNNAN	pancoea ana
5	795.5	39.8	382	Q47844	pancoea ag9
6	783.5	39.2	382	O8GCS2	pancoea ag9
7	782.5	39.1	386	O8VUJ6	pancoea ag9
8	779.5	39.0	385	O93C18	xanthobacte
9	766.5	38.3	394	O9K1X3	bradyrhizob
10	750	37.5	386	CRTY_ERWHE	erwinia her
11	671	33.6	391	Q7M2F1	photobact
12	648.5	32.4	379	O6MMA6	bdellovibri
13	551.5	27.6	434	O06756	erythrobact
14	195.5	9.8	501	LCYB_ARATH	citrus sine
15	191.5	9.6	503	CCS_CITSI	citrus sine
16	187	9.4	490	O84V99	zea mayi (m
17	184.5	9.2	489	O6YUJ2	oryza sativ
18	178.5	8.9	411	O83WZ9	strepomyce
19	175.5	8.8	504	O64HC6	citrus sine
20	174.5	8.7	397	O82PA6	strepomyce
21	174.5	8.7	448	O93HP4	strepomyce
22	174.5	8.7	504	O8LPP7	citrus sine
23	174.5	8.7	524	O8GTR2	citrus unsh
24	174.5	8.7	504	O9XGX3	citrus para
25	173.5	8.7	504	O6XO45	citrus maxi
26	173	8.7	418	P72451	strepomyce
27	172.5	8.6	504	O9MS46	citrus sine
28	170.5	8.5	498	LCYB_CAPAN	citrus sine
29	170.5	8.5	504	O66NC9	citrus sine
30	170.5	8.5	504	O66ND0	citrus sine
31	164.5	8.2	525	O761W9	cryptomeria

32	164.5	8.2	525	2	O761Z0	cryptomeria
33	164.5	8.2	525	2	O761Z6	cryptomeria
34	164.5	8.2	525	2	O761J3	cryptomeria
35	164.5	8.2	525	2	O761J4	cryptomeria
36	164.5	8.2	525	2	O761J8	cryptomeria
37	164	8.2	525	2	O761J5	cryptomeria
38	163.5	8.2	502	2	O9AXL1	strepomyce
39	163.5	8.2	525	2	O761R1	taxodium pala
40	163.5	8.2	525	2	O761R4	taxodium di
41	162	8.1	410	1	Y801_DEIRA	cryptomeria
42	161.5	8.1	500	1	LCYB_TOBAC	delnoccous
43	161	8.1	500	1	LCYB_TOBAC	lycopersico
44	159	8.0	498	2	O9M424	nicotiana t
45	157.5	7.9	430	2	O818H5	solanum tub

ALIGNMENTS

RESULT 1	PRELIMINARY	PRT	382 AA.
ID P94791			
AC P94791			
DT 01-MAY-1997 (TREMBLrel. 03, Created)			
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)			
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE Lycopene cyclase.			
GN Name=crly;			
OS Flavobacterium sp. ATCC 21568.			
OC Bacteria; Bacteroidetes; Flavobacteriales;			
OC Flavobacteriaceae; Flavobacterium.			
OX NCBI_TaxID=50286;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=R1534;			
RX MEDLINE=97186694; PubMed=9034310; DOI=10.1016/S0378-1119(96)00624-5;			
RA Pasamontes L., Hug D., Tessier M., Hohmann H.P., Schierle J.,			
van Loon A.P.;			
RT "Isolation and characterization of the carotenoid biosynthesis genes			
of Flavobacterium sp. strain R1534.";			
RL Gene 185:35-41(1997).			
DR EMBL; U62808; AAC44851.1; "			
DR GO; GO:0045436; F:lycopene beta cyclase activity; IEA.			
DR GO; GO:0016705; F:oxidoreductase activity, acting on paired d. . .; IEA.			
DR GO; GO:0016117; P:carotenoid biosynthesis; IEA.			
DR InterPro; IPR010108; Carotene_cycl.			
DR Pfam; PF05897; Crly; 1.			
DR TIGRFAMs; TIGR01790; carotene_cycl; 1.			
DR TIGRFAMs; TIGR01789; lycopene_cycl; 1.			
SQ SEQUENCE 382 AA; 42369 MM; 6DBS452F45D9EF4B CRC64;			
Query Match	100.0%;	Score 1999;	DB 2; Length 382;
Best Local Similarity	100.0%;	Pred. No. 1.1e-145;	
Matches 382; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY 1 MSHDLLIAGAGLGGALIALAVDRRPPDARIWLDNRSGSDQHTWSCHPDTLSPEHLAL 60			
DB 1 MSHDLLIAGAGLGGALIALAVDRRPPDARIWLDNRSGSDQHTWSCHPDTLSPEHLAL 60			
QY 61 SIRRGEWTDQVAPPDHSRRLLTGYGSTEAGALIGLQGVLDLRNTHVATLDDPGATLT 120			
DB 61 SIRRGEWTDQVAPPDHSRRLLTGYGSTEAGALIGLQGVLDLRNTHVATLDDPGATLT 120			
QY 121 DGSRIEACVIDARGAVETPHLTVGFQKVGVEITDAPHGVERPMIMDATVPQMDGYRF 180			
DB 121 DGSRIEACVIDARGAVETPHLTVGFQKVGVEITDAPHGVERPMIMDATVPQMDGYRF 180			
QY 181 IYLLPSPFRLIETRTYSDGDLDDGALAAQSLIYAARGTGGEMRERERILPALAH 240			
DB 181 IYLLPSPFRLIETRTYSDGDLDDGALAAQSLIYAARGTGGEMRERERILPALAH 240			
QY 241 DALGFWRBDAQAVPVGAGLGFHPVTGYSLPYAQVADALIAARDLTTASARAVRGNAI 300			

```

Db      241 DAIGWRBHAQGA VVGAGLGFHPTVGS LSLPYAAQVDAIARLDLTASARARAGMAI 300
Qy      301 DRADDRFLRLINMLFRGCPDPRRYLLORFYRLPOPLIERFYAGRLTLADRLITVGR 360
Db      301 DRADDRFLRLINMLFRGCPDPRRYLLORFYRLPOPLIERFYAGRLTLADRLITVGR 360
Qy      361 PEIPLSOAVRCLPERPILQERA 382
Db      361 PEIPLSOAVRCLPERPILQERA 382

```

RESULT 2

```

ID      CTRY AGRU      STANDARD;      PRT;      386 AA.
AC      P54974;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Lycopene cyclase.
GN      Name=crly;
OS      Agrobacterium aurantiacum.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OX      Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
RN      NCBI_TaxID=41155;
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96062243; PubMed=7592436;
RA      Misawa N., Satomi Y., Kondo K., Yokoyama A., Kajiwara S., Saito T.,
RA      Ohnami T., Miki M.;
RT      "Structure and functional analysis of a marine bacterial carotenoid
RT      biosynthesis gene cluster and astaxanthin biosynthetic pathway
RT      proposed at the gene level."
RL      J. Bacteriol. 177:6575-6584(1995).
CC      -!- FUNCTION: Catalyzes the cyclization reaction which converts
CC      lycopene to beta-carotene.
CC      -!- PATHWAY: Carotenoid biosynthesis. Involved in astaxanthin
CC      biosynthetic pathway.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; D58420; BAA09593.1; -.
DR      InterPro; IPR010108; Carotene_cycl.
DR      Pfam; PF05897; Crly; 1.
DR      TIGRFAMs; TIGR01790; carotene-cycl; 1.
DR      TIGRFAMs; TIGR01789; lycopene_cycl; 1.
DR      Carotenoid biosynthesis.
SQ      SEQUENCE 386 AA; 42202 MW; E14C41B688AFT78AC CRC64;

```

Query Match 72.8%; Score 1456; DB 1; Length 386;
 Best Local Similarity 71.8%; Pred. No. 7.3e-104;
 Matches 277; Conservative 39; Mismatches 66; Indels 4; Gaps 2;

```

Qy      1 MSHDLLIAGAGLIALAVDRRDPARIWMLDARSGPSDOHTWSCDPTDLSPEMLARL 60-
Db      1 MTHDVLLAGAGLIALALRAARPDRLVLLDHAAGPSDGHWTSCDPTDLSPEMLARL 60
Qy      61 SPIRGGEWTDQVAPPDHSRLTTGYGSLIEGALIGLL--QGVDLRWNTHTVATLDDTGAT 118
Db      61 KPLRRANMPDQEVREPRHARRLATGYSLDGAALADAVRSGAEIRWMSDIALDLDAQAT 120
Qy      119 LTDGRIEAAVCYIDARGA VETPHLTGFOKFGVGEIETDAPHGVRPMMIMATVPQMDGY 178
Db      121 LSCGRIEAGVLDGRGAQPSRLITVGFOKFGVGEIETDPRPGVRPMMIMATVPQMDGY 180
Qy      179 RFIYLLPSPSTRILLIEDTRYSDGDLDDGALAQAASLDYVARRGWTGOEMRBRGILPITAL 238

```

```

Db      181 RFIYLLPSPSTRILLIEDTRYSDGDLDDGALAQAASLDYVARRGWTGOEMRBRGILPITAL 240
Qy      239 AHDAIGFRRDHAQGA VVGAGLGFHPTVGS LSLPYAAQVDAIARLDLTASARARAVR 296
Db      241 AHDAIGFRRDHAQGA VVGAGLGFHPTVGS LSLPYAAQVDAIARLDLTASARARAVR 300
Qy      297 GWAIDRADRDFRLINMLFRGCPDPRRYLLORFYRLPOPLIERFYAGRLTLADRLIRI 356
Db      301 DVAIDRADRDFRLINMLFRGCPDPRRYLLORFYRLPOPLIERFYAGRLTLADRLIRI 360
Qy      357 VTGRPIPLSOAVRCLPERPILQERA 382
Db      361 VTGRPIPLSOAVRCLPERPILQERA 382

```

RESULT 3

```

ID      CTRY AGRU      PRELIMINARY;      PRT;      386 AA.
AC      Q9RLH5;
DT      01-MAY-2000 (TRENDELrel. 13, Created)
DT      01-MAY-2000 (TRENDELrel. 13, Last sequence update)
DT      01-MAR-2004 (TRENDELrel. 26, Last annotation update)
DE      Lycopene beta-cyclase.
GN      Name=crly;
OS      Paracoccus marcusii.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OX      Rhodobacteraceae; Paracoccus.
RN      NCBI_TaxID=59779;
RP      SEQUENCE FROM N.A.
RX      STRAIN=MH1;
RA      Harker M., Hirschberg J.;
RT      Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
RL      EMBL: Y15112; CAB56061.1; -.
DR      GO; GO:0045436; F:lycopene beta cyclase activity; IEA.
DR      GO; GO:0016705; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR      GO; GO:0016117; P:carotenoid biosynthesis; IEA.
DR      InterPro; IPR010108; Carotene_cycl.
DR      Pfam; PF05897; Crly; 1.
DR      TIGRFAMs; TIGR01790; carotene-cycl; 1.
DR      TIGRFAMs; TIGR01789; lycopene_cycl; 1.
SQ      SEQUENCE 386 AA; 42283 MW; 91952F0F3F3F6800 CRC64;

```

Query Match 70.3%; Score 1405; DB 2; Length 386;
 Best Local Similarity 69.2%; Pred. No. 6.2e-100;
 Matches 267; Conservative 43; Mismatches 72; Indels 4; Gaps 2;

```

Qy      1 MSHDLLIAGAGLIALAVDRRDPARIWMLDARSGPSDOHTWSCDPTDLSPEMLARL 60
Db      1 MTHDVLLAGAGLIALALRAARPDRLVLLDHAAGPSDGHWTSCDPTDLSPEMLARL 60
Qy      61 SPIRGGEWTDQVAPPDHSRLTTGYGSLIEGALIGLL--QGVDLRWNTHTVATLDDTGAT 118
Db      61 KPLRRANMPDQEVREPRHARRLATGYSLDGAALADAVRSGAEIRWMSDIALDLDAQAT 120
Qy      119 LTDGRIEAAVCYIDARGA VETPHLTGFOKFGVGEIETDAPHGVRPMMIMATVPQMDGY 178
Db      121 LSCGRIEAGVLDGRGAQPSRLITVGFOKFGVGEIETDPRPGVRPMMIMATVPQMDGY 180
Qy      179 RFIYLLPSPSTRILLIEDTRYSDGDLDDGALAQAASLDYVARRGWTGOEMRBRGILPITAL 238
Db      181 GFIYLLPSPSTRILLIEDTRYSDGDLDDGALAQAASLDYVARRGWTGOEMRBRGILPITAL 240
Qy      239 AHDAIGFRRDHAQGA VVGAGLGFHPTVGS LSLPYAAQVDAIARLDLTASARARAVR 296
Db      241 AHDAIGFRRDHAQGA VVGAGLGFHPTVGS LSLPYAAQVDAIARLDLTASARARAVR 300
Qy      297 GWAIDRADRDFRLINMLFRGCPDPRRYLLORFYRLPOPLIERFYAGRLTLADRLIRI 356
Db      301 DVAIDRADRDFRLINMLFRGCPDPRRYLLORFYRLPOPLIERFYAGRLTLADRLIRI 360

```



```

Db          363  SGKPPVLAALQAI 377

RESULT 6
O8GCS2      PRELIMINARY;      PRT;      382 AA.
ID   O8GCS2      PRELIMINARY;      PRT;      382 AA.
AC   O8GCS2;
DT   01-MAR-2003 (TREMBLrel. 23, Created)
DT   01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE   Lycopene beta-cyclase.
GN   Name=ctry;
OS   Pantoea stewartii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Pantoea.
OX   NCBI_TaxID=66269;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 8200.
RA   desouza M.L., Kollmann S.R., Schroeder W.A.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AY166713; AAN85598.1; -.
DR   GO; GO:0045436; F:lycopene beta cyclase activity; IEA.
DR   GO; GO:0016705; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR   GO; GO:0016117; F:carotenoid biosynthesis; IEA.
DR   InterPro; IPR010108; Carotene_cycl.
DR   Pfam; PF05897; Ctry; 1.
DR   TIGRFAMs; TIGR01790; carotene-cycl; 1.
DR   TIGRFAMs; TIGR01789; lycopene_cycl; 1.
SQ   SEQUENCE 382 AA; 43257 MW; 7893D4976C5FCCB1 CRC64;

Query Match          39.2%; Score 783.5; DB 2; Length 382;
Best Local Similarity 43.7%; Pred. No. 4.3e-52;
Matches 165; Conservative 64; Mismatches 136; Indels 13; Gaps 6;

OY   3 HDLIAGAGLGGALITAVRRRPPARIYVMDARSGPSDQHTWSDCHTDLS---EWLAR 59
DB   5 YDLIVGAGLANGLIALRLQOQHPMRILLIEAGPEAGGNHTWSFHEEDLTINQHRMTA- 63
OY   60 LSPIRGEMTQEVAFPDHSRRLTTGYGSIKAGLIGLIQ---GVDLKNTHVATLDDTG 116
DB   64 --PLVVHHPDYOVRFPQRRHNSGYCVTSRHPAGILRQOPGHLMLHTVSAVHAES 121
OY   117 ATLTDGSRIEACVYDARGAVETPHLTGFKFQVEIETDAPHGVERPMIMDATVPOMD 176
DB   122 VQLAGRIIHAISTVIDRGYTDSALRVGFQAFIQEWQLSPHGLSEPIIMDAIVDOQN 181
OY   177 GYRFIYLLPFSPTRLIEDTRYSDGGLDDGALAQASLDVYARRWGTQE--NRERGITLP 235
DB   182 GYRFYVYTLPLSATLLIEDTHYIDKANLQAEERARONIRDYARQGWPLQTLIREQGLP 241
OY   236 IALAHDAIGFMRDHAOGAVPVGLGAGLFHPVTGYSLPAAQVADIAARDL--TTSASARA 294
DB   242 ITLVGDNRQFQWQOQPOAC--SGLRAGLFHPVTGYSLPVALADRLSADVTSVHOT 299
OY   295 VVGMAIDRADRDRLFLRLINRMLFRGCPDPRRYRLIQRFYRLPOPLIERFYAGRLTLADRL 354
DB   300 IAHFAGQWQOQGFRRMLNRMFLFLAGPASPASRWVQRFGLPEDLIAFYAGKLTVTBRL 359
OY   355 RIIVGRPPILPSQAVRCL 372
DB   360 RLISGKPPVPVLAALQAI 377

RESULT 7
O8VU06      PRELIMINARY;      PRT;      386 AA.
ID   O8VU06      PRELIMINARY;      PRT;      386 AA.
AC   O8VU06;
DT   01-MAR-2002 (TREMBLrel. 20, Created)
DT   01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE   Ctry protein.

```

```

GN   Name=ctry;
OS   Pantoea agglomerans pv. melleiae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Pantoea.
OX   NCBI_TaxID=182454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   Kamuntzen H., Hirata R.;
RA   Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AB076662; BAB79602.1; -.
DR   GO; GO:0045436; F:lycopene beta cyclase activity; IEA.
DR   GO; GO:0016705; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR   GO; GO:0016117; F:carotenoid biosynthesis; IEA.
DR   Pfam; PF05897; Ctry; 1.
DR   TIGRFAMs; TIGR01790; carotene-cycl; 1.
DR   TIGRFAMs; TIGR01789; lycopene_cycl; 1.
SQ   SEQUENCE 386 AA; 43674 MW; 1E7DB42C08447AE7 CRC64;

Query Match          39.1%; Score 782.5; DB 2; Length 386;
Best Local Similarity 43.8%; Pred. No. 5.2e-52;
Matches 168; Conservative 62; Mismatches 137; Indels 17; Gaps 8;

OY   3 HDLIAGAGLGGALITAVRRRPPARIYVMDARSGPSDQHTWSDCHTDLS---EWLAR 59
DB   4 YDLIVGAGLANGLIALRLQOQPSRLILLIDAEEREPANHWSFHAEEDLTETQHRMTA- 62
OY   60 LSPIRGEMTQEVAFPDHSRRLTTGYGSIKAGLIGLIQ---GVDLKNTHVATLDDTG 116
DB   63 --PLVVHHPDYOVRFPQRRHNSGYCVTSRHPAGILRQOPGHLMLHTVSAVHAES 120
OY   117 ATLTDGSRIEACVYDARGAVETPHLTGFKFQVEIETDAPHGVERPMIMDATVPOMD 176
DB   121 VTLDDGRLTEADAVIDGKGYPDGLRMGFQSFQGEWQLSEPHLTAPIIMDAIVDOQA 180
OY   177 GYRFIYLLPFSPTRLIEDTRYSDGGLDDGALAQASLDVYARRW--TGQEMRRRGITLP 235
DB   181 GYRFYVYTLPLSATLLIEDTHYIDNATLIEGRARONIDVYARQGWPLQTLIREQGLP 240
OY   236 IALAHDAIGFMRDHAOGAVPV--GLGAGLFHPVTGYSLPAAQVADIA--ARDLTTASARR 293
DB   241 ITLVGDVAAPQNR---DLPCSLRAGLFHHTTGSLSPLAVALADRLAQMTFTSETLHA 297
OY   294 AVRGMAIDRADRDRLFLRLINRMLFRGCPDPRRYRLIQRFYRLPOPLIERFYAGRLTLADR 353
DB   298 TIQGFASQAWQOQGFRRMLNRMFLFLAGPADQORWQMRFGYGLPEGLIARFYAGKLTLPDR 357
OY   354 RIIVGRPPILPSQAVRCL--PER 375
DB   358 RLISGKPPVPVLAALQAIPTPHR 381

RESULT 8
O93CT8      PRELIMINARY;      PRT;      385 AA.
ID   O93CT8      PRELIMINARY;      PRT;      385 AA.
AC   O93CT8;
DT   01-DEC-2001 (TREMBLrel. 19, Created)
DT   01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE   Lycopene cyclase.
GN   Name=ctry;
OS   Xanthobacter sp. (strain Py2).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Hyphomicrobiaceae; Xanthobacter.
OX   NCBI_TaxID=78245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Py2;
RA   MEDLINE=22176293; PubMed=12189420;
RA   Larsen R.A., Wilson M.M., Guse A.M., Metcalf W.W.;
RT   "Genetic analysis of pigment biosynthesis in Xanthobacter
autotrophicus Py2 using a new, highly efficient transposon mutagenesis
system that is functional in a wide variety of bacteria.";
RL   Arch. Microbiol. 178:193-201 (2002).

```


[illegible]

or send an email to license@sb-sib.ch).

CC EMBL; M87280; AAA64980.1; --
 DR PIR; S52981; S52981.
 DR InterPro; IPR010108; Carotene_cycl.
 DR InterPro; IPR008461; CrtY.
 DR Pfam; PF05897; CrtY; 1.
 DR TIGRfams; TIGR01790; carotene_cycl; 1.
 DR TIGRfams; TIGR01789; lycopen_cycl; 1.
 KW Carotenoid biosynthesis.
 SQ SEQUENCE 366 AA; 43341 MW; F4A40563BFCFA980 CRC64;

Query Match 37.5%; Score 750; DB 1; Length 386;

Best Local Similarity 43.0%; Pred. No. 1.7e-49;

Matches 166; Conservative 51; Mismatches 149; Indels 20; Gaps 6;

QY 4 DLIAGAGLGGALTAALVRRDRPDARIYMLDARGSPSOHTWSCHDTLSPWMLARLSP 63
 DB 3 DLIIVGGGLANGLIAMRLRQRYPOINLLITEGEPGNGHTWSFHEDLTPGOHMLAPL 62
 QY 64 RRGWTDQEVAFPDHSRLTTGSGIEAGALIGLQ---GVDLRNNTVATLDTGATLT 120
 DB 63 VAHMPGVQVQPDRLRRRLARGYSITSRFPAALHQAIGENIWLNCVSIVLPRSVRLA 122
 QY 121 DGSRIEACVIDARGAVETPHLTGFGKFGVGEIETDAPHYERPMINDATVPQMDGYR 180
 DB 123 NGEALLAGAVIDRGVLTASAMQGYQLFLGQOMRLTDPHGLTVPLMDATVAQGGYRF 182
 QY 181 IYLLPFSPRLILIEDTRYSDGDLDDGALAQSILVYARRGW-TGQEMRRERGLPIALA 239
 DB 183 VYTLPLSADTLIEDTRYANVPORDNALRQTVTVYVHASKWQLAQLEEEETGLPITLA 242
 QY 240 HDAIFWMDHQAQAVPVGLAGLFHPVTGSLPYAAOVADALIA---ARDLTASARR 293
 DB 243 GDIOQLMAD-AGVRSRSGRAGLHFPTTGYSLPLVALADALADSPRLGSVLYLTQCF 301
 QY 294 AVRGALIDRADDRFLRLNRLFRGCPDRRYRLLORFYRLPOPLIEFYAGRLTLADR 353
 DB 302 AERHNR-----RQGFRLNRLFLAGRENRMRVMQRFYGLPEYVERFYAGRLSLPDK 356
 QY 354 LRIYVGRPIPLSQAVRC---LPRR 375
 DB 357 ARLTGKPPVPLGEAMRAALNHPDR 382

RESULT 11

Q7MZFL PRELIMINARY; PRT; 391 AA.

AC Q7MZFL;
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE CrtY protein.
 GN Name=CrtY; Ordered locus names=plu4341;
 OS Photorhabdus luminescens (subsp. launondii);
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Photorhabdus.
 NC NCBI_TaxID=141679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TT01;
 RX MEDLINE=22957627; PubMed=14528314;
 RA Duchaud E., Ruenstok C., Frangeul L., Buchrieser C., Givaudan A.,
 RA Taourit S., Bocs S., Boursaux-Bude C., Chandler M., Charles J.-F.,
 RA Dassa E., Derose R., Derzelle S., Freyssiuet G., Gaudriault S.,
 RA Medigue C., Lanois A., Powell K., Siglier P., Vincent R., Wingate V.,
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
 RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
 luminescens".
 RL Nat. Biotechnol. 21:1307-1313(2003).
 DR EMBL; BX571873; CAE16713.1; --
 DR Photolab; plu4341; --
 DR GO; GO:0045436; F:lycopene beta cyclase activity; IEA.

DR GO; GO:0016705; F:oxidoreductase activity, acting on paired d. .; IEA.
 DR GO; GO:0016117; P:carotenoid biosynthesis; IEA.
 DR InterPro; IPR010108; Carotene_cycl.
 DR InterPro; IPR008461; CrtY.
 DR Pfam; PF05897; CrtY; 1.
 DR TIGRfams; TIGR01790; carotene_cycl; 1.
 DR TIGRfams; TIGR01789; lycopen_cycl; 1.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 44725 MW; 240ECC992E10B21F CRC64;

Query Match 33.6%; Score 671; DB 2; Length 391;

Best Local Similarity 41.0%; Pred. No. 2e-43; Indels 14; Gaps 8;

Matches 155; Conservative 57; Mismatches 152;

QY 4 DLIAGAGLGGALTAALVRRDRPDARIYMLDARGSPSOHTWSCHDTLSP---EWLRL 60
 DB 6 DLIIVGGGLANGLIAMRQCKPHRLVLLINTERIGNHTWSFHQDLTEAHEHMLAPL 65
 QY 61 SPIRGWTDQEVAFPDHSRLTTGSGIEAGALIGLQ---GVDLRNNTVATLDTGCA 117
 DB 66 IYVR---WSGYDVIFPAFQRTLPHSYFISITQHFAHIAVIGERIQRLVQELTPQKV 122
 QY 118 TLTGSRIEACVIDARGAVETPHLTGFGKFGVGEIETDAPHYERPMINDATVPQMDG 177
 DB 123 YLQDSSLSAGAVIDRGMRPGPTGSGTQAFQGEWLESHSLTHPLMDTSVGQDTG 182
 QY 178 YRFIYLPFSPRLILIEDTRYSDGDLDDGALAQSL-DYARRGW-TGQEMRRERGLP 235
 DB 183 YRFIYLPFSSRLILIEDTHVDRGP-PDKLSQATIAEVAKKGWKGLIRESSGLP 241
 QY 236 IALAHADIGFWRDHAQAVPVGLAGLFHPVTGSLPYAAOVADALIA-RDLTASARRA 294
 DB 242 ITLTDFTSFQAQLA-GQPTGRLAALFHPPTGSLPFAILARIVLPELTDTSFLT 300
 QY 295 VRGALIDRADDRFLRLNRLFRGCPDRRYRLLORFYRLPOPLIEFYAGRLTLADR 354
 DB 301 LRDYRQGMQORFRLNRLFLAGDPQGMQWRFYLSPLIARFYALQNSVKA 360
 QY 355 RIYVGRPIPLSQAVRCL 372
 DB 361 RILICKPPVPIKGLAKAM 378

RESULT 12

O6MMA6 PRELIMINARY; PRT; 379 AA.

AC O6MMA6;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Lycopene cyclase.
 GN Ordered locus names=Bd1730;
 OS Bdellovibrio bacteriovorus.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
 OC Bacteriia; Bacteriia; Bdellovibrio.
 NC NCBI_TaxID=959;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
 RX PubMed=14752164; DOI=10.1126/science.1093027;
 RA Rendulic S., Jagtap P., Rosinus A., Bpinger M., Baar C., Lanz C.,
 RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
 RA Sockett R.E., Schuster S.C.;
 RT "A predator umasked: 11fe cycle of Bdellovibrio bacteriovorus from a
 genomic perspective".
 RL Science 303:689-692(2004).
 DR EMBL; BX842650; CAE79599.1; --
 DR GO; GO:0045436; F:lycopene beta cyclase activity; IEA.
 DR GO; GO:0016705; F:oxidoreductase activity, acting on paired d. .; IEA.
 DR GO; GO:0016117; P:carotenoid biosynthesis; IEA.
 DR InterPro; IPR010108; Carotene_cycl.
 DR InterPro; IPR008461; CrtY.
 DR Pfam; PF05897; CrtY; 1.

DR TIGRFAMS: TIGR01790; carotene_cycl, 1.
 DR TIGRFAMS: TIGR01789; lycopene_cycl, 1.
 KW Complete proteome.
 SQ SEQUENCE 379 AA; 42420 MW; 1A3F7D98A512A54 CRC64;

Query Match 32.4%; Score 648.5; DB 2; Length 379;
 Best Local Similarity 37.4%; Pred. No. 1.1e-41;
 Matches 144; Conservative 70; Mismatches 138; Indels 33; Gaps 9;

QY 4 DLIAGAGSGALIALAVDRRDPARIWMDARSGSDQHTWSCHDTLSPF--WLAKRS 61
 DB 11 DCLIVGGGAGGILLQALRTEQPELVKLLERGTQAGNHTWSFHGSDVPGATWQSLI 70
 QY 62 PIRGEMTQEVAFPHDSRLTTGYSIEA---GALIGLQGVDRMTWHTVATLDDTGA 117
 DB 71 ---SKTPAYEVRFPYQKRIQSSYSIKADPHQKLLG-QHSGQILLQASVQEVARDV 126
 QY 118 TLTDGSRIBACVVDARGAVETPHLTVGQKVEVEIETDAPHGERPMIMATVPQMDG 177
 DB 127 TLDDGMFPAKCVIDARGGAADVARGYQKFGVLVKSQPHGLNHVILKDLVLPQVDG 186
 QY 178 YRFYTLPPSPTRILIEDTRYSDGDLDDGALAQAASLDYARRGW-TGQEMREREGILPI 236
 DB 187 YRFVYILPMSETELLVEDTYVSTPDLDTALSKGILDYIAGKWTESVIRQEVGCLPL 246
 QY 237 ALAHDAIGFWRDHAQAVPVGGA--GLPHPTVGSLLPYAAQVADAIAARDLTASARRA 294
 DB 247 DLYDVAV---EGSGSGPPLNAGASGVQPVGTGYFPQTAVCAVLAASSLDT----- 296
 QY 295 VEGWA-----IDRADRDRLFLNLNMLFRGCPDRRYLLQRFYRLPQPLIERFYAGR 347
 DB 297 ---WGVVLPAQLQNTYMKQARLYRLINRMFLAAVPEKRYVILERFLLEBALIERFYQGR 353
 QY 348 LFLADRLRIVTGRPPILPSQAVRCL 372
 DB 354 LTVLDQVRILCGKPPVSVWRALKSL 378

RESULT 13

006756 PRELIMINARY; PRT; 434 AA.
 AC 006756;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Lycopene cyclase.
 GN Name=crty;
 OS Erythrobacter longus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
 OC Sphingomonadaceae; Erythrobacter.
 OK NCBI_TaxId=1044;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=och 101;
 RX MEDLINE=97311406; PubMed=9168123; DOI=10.1016/S0378-1119(96)00788-3;
 RA Matsumura H., Takayama H., Hasekabe E., Burgess J.G., Matsumura T.,
 RT "Cloning, sequencing and expression of the carotenoid biosynthetic
 RT lycopene cyclase and phytoene desaturase, from the aerobic
 RT photosynthetic bacterium Erythrobacter longus sp. strain Och 101 in
 RT Escherichia coli.";
 RT Gene 189:169-174(1997).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=och 101;
 RA matsumura h., hasekabe e., matsumura t.,
 RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: D83513; BAA20275.1; -;
 DR GO: GO:0045436; F:lycopene beta cyclase activity; IEA.
 DR GO: GO:0016705; P:oxidoreductase activity; acting on paired d. . .; IEA.
 DR GO: GO:0016117; P:carotenoid biosynthesis; IEA.
 DR InterPro: IPR010108; Carotene_cycl.
 DR InterPro: IPR008461; Crty.
 DR Pfam: PF05897; Crty; 1.

DR TIGRFAMS: TIGR01790; carotene_cycl, 1.
 DR TIGRFAMS: TIGR01789; lycopene_cycl, 1.
 KW Complete proteome.
 SQ SEQUENCE 434 AA; 48018 MW; 9E440E09BA5AEF47 CRC64;

Query Match 27.6%; Score 551.5; DB 2; Length 434;
 Best Local Similarity 36.0%; Pred. No. 3.7e-34;
 Matches 142; Conservative 55; Mismatches 164; Indels 33; Gaps 10;

QY 2 SHDLIAGAGSGALIALAVDRRDPARIWMDARSGSDQHTWSCHDTLSPF--WLAKRS 61
 DB 14 SCDCALVGGGAGGILLQALRTEQPELVKLLERGTQAGNHTWSFHGSDVPGATWQSLI 73
 QY 62 PIRGEMTQEVAFPHDSRLTTGYSIEA---GALIGLQGVDRMTWHTVATLDDTGA 115
 DB 74 DFRQIDMEGVEYRFPKRYKLTAVRSWASDHFHGLRALPEGSVILGRKAVG-LDAR 132
 QY 116 GATLT---DGSRIEACVVDARGAVETPHLTVGQKVEVEIETDAPHGERPMIMATVP 170
 DB 133 GVDLAPSQYGPATRIARSVIDCRSEKPSAHKGGQVFLGHRMRLOEBHGVENPIVMDA 192
 QY 171 TVPQW-----DGYRITVLLPSPTRILIEDTRYSDGDLDDGALAQAASLDYARRGW-T 223
 DB 193 TVDQALPHNGSGSYRPPVYLPGSHDVFLIEDTYADPDLDDNVALSGRIDYARANGWEN 252
 QY 224 GQEMREREGILPIALAHDAIGFWRDHAQ--GAVPGLAGLPHPTVGSLLPYAAQVADAIA 281
 DB 253 GTPVHHEAGVLPVLTGSD-FSAVQDEVRIIPGVALIAGARGFTPHPLSTYMCVAVENALAM 311
 QY 282 A-----ARDLTASARRAVRGMAIDRADRDRLFLNLNMLFRGCPDRRYLLQRFYRL 335
 DB 312 ASDPDLGSGQLAALFSDSRARRHMS-----KTGYRLLALFLFAKPEKRVKVFQRFYGL 366
 QY 336 PQLITERFYAGRLTLADRLRIVTGRPPILPSQAV 369
 DB 367 RSGILIERFYAARSNTFDKRVLMGEPVAIHSAI 400

RESULT 14

LCYB_ARATH STANDARD; PRT; 501 AA.
 ID LCYB_ARATH
 AC Q38933; Q39145;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Lycopene beta cyclase, chloroplast precursor (EC 1.14.-.-).
 GN Name=LCY1; Synonyms=LYC; OrderedLocustNames=Ac3g10230;
 ORFNames=F14P13.17;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsie.
 OK NCBI_TaxId=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=96434545; PubMed=8837512;
 RA Cunningham F.X., Jr., Fosson B., Sun Z., McDonald K.A., Dellapenna D.,
 RA Gantt E.;
 RT "Functional analysis of the beta and epsilon lycopene cyclase enzymes
 RT of Arabidopsis reveals a mechanism for control of cyclic carotenoid
 RT formation.";
 RT Plant Cell 8:1613-1626(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Wassilewskija;
 RA Scolnik P.A., Bartley G.B.;
 RT "Nucleotide sequence of lycopene cyclase from Arabidopsis.";
 RL (er) Plant Gene Register PGR95-019.
 RN (3)
 RP SEQUENCE FROM N.A.
 RA Giuliano G., Rosati C., Santangelo G.;
 RT "Gene structure and regulation of the carotenoid biosynthesis pathway
 RT in Arabidopsis thaliana.";

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cnv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713; DOI=10.1038/35048706;
 RA Salanoubat M., Lemcke K., Rieger M., Ansgorge W., Unseld M.,
 RA Salanoubat B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
 RA Delaney M., Boutry M., Griwell L.A., Maché R., Puligdomench P.,
 RA De Simone V., Choisme N., Artiguenave F., Robert C., Broctier P.,
 RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Queller F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Bens V.,
 RA Muehlbach E., Drzozek H., Ertle H., Jordan N., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,
 RA Conrad A., Hohnischer K., Kauer G., Loehner T.-H., Nordstiek G.,
 RA Reichelt J., Schärfe M., Schoen O., Barques M., Terol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottewillder B., Duchemin D.,
 RA Cooke R., Laudie M., Berger-Lilauro C., Purnelle B., Masuy D.,
 RA de Haan M., Maare A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argitrou A., Flores M., Iignori R., Vitale D.,
 RA Mannheim G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts R., Wu D., Peterson J., Van Aken S.,
 RA Creasy T.H., Haas B., Malt R., Gall J.E., Feldblyum T.V.,
 RA Pai G., Miltcher J., Sellers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Ventner J.C.,
 RA Frazer C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Idehara K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsuno M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shingo S., Takeuchi C., Wada T.,
 RA Matenabe A., Yamada M., Yasuda M., Tabata S.;
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:820-822(2000).
 CC -!- FUNCTION: Catalyzes the double cyclization reaction which converts
 CC lycopene to beta-carotene and neurosporene to beta-zeacarotene.
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: Belongs to the lycopene cyclase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 CC EMBL, U50739; AAB53337.1; -;
 CC EMBL, L40176; AAB81880.1; -;
 CC EMBL, AF117256; AAF82388.1; -;
 CC EMBL, AC009400; AAF02819.1; -;
 CC InterPro: IPR010108; Carotene_cycl.
 CC InterPro: IPR008671; Lycopene_cycl.
 CC InterPro: IPR000205; NAD_BS.
 CC Pfam: PF05834; Lycopene_cycl; 1.
 CC TIGFAMS: TIGR01790; carotene_cycl; 1.
 CC Carotenoid biosynthesis; Chloroplast; NAD; Oxidoreductase;
 CC Transist peptide.
 CC
 CC TRANSIT 1 80 Chloroplast (Potential).
 CC CHAIN 81 501 Lycopene beta cyclase.
 CC NP_BIND 85 113 NAD (Potential).
 CC FT CONFLICT 31 31 NAD (Potential).
 CC FT CONFLICT 243 243 V -> I (in Ref. 2).
 CC FT CONFLICT 243 243 V -> I (in Ref. 2).
 CC SEQUENCE 501 AA; 56176 MW; C30145780DBDC4E2 CRC64;
 CC
 CC Query Match 9.8%; Score 195.5; DB 1; Length 501;
 CC Best Local Similarity 23.4%; Pred. No. 1.1e-06;
 CC Matches 111; Conservative 66; Mismatches 144; Indels 153; Gaps 27;
 CC
 CC 4 DLIAGAGISGALTLALVDRPRDARIWLDARSGPSDQHTSCSDTDLSPFWLARLSTI 63
 CC DB 84 DLAIVGAGPAGLAVAGQVSE--AGLSVCSID--PSPRLTW-----PN 121

QY 64 RRGWTDQ-----EVAEPDH--SRRLTGYGSIAGALIG-----LLQ 99
 DB 122 NYGVAVDEFEAMDLLDCLDTTWSGAVVYVDEGVKKDLSRPGYRVNRKQLKSKMLQKCTTN 181
 QY 100 GVDLAKM--TIVATLDTGAL--TDGSRIFAACYIDARG-----AVETHLVTPQKF 149
 DB 182 GVKFHQSRYTVNVH--EENSTVSCSDGVKIASVLDATGFSRCLVQYDKY--NPGYVA 239
 QY 150 VGVETETA--PHGVRPIMATVPQMDGY-----RFYLLPSPFRLIEDTR 197
 DB 240 YGIVAEVDHGFPDDKVPMDWRKHDLSYELKERNSKIPTFLYAMFSSNRIFLEET- 298
 QY 198 YSDGDDLDGALQASL-----DYARRCWYQEMRR-----ERGIPLALADALIGFW 246
 DB 299 -----SLVARGLMEDIQEMMARLKLHGINVGRIBEDRCVJPMG----- 340
 QY 247 RDHAGANP-----VIGG--AGLFHPYTG-----SLPYAQQVADALARLTASARRA 294
 DB 341 -----GPLPVLPRVVGIGTAGWVHPSTGYVARTTAAAPIVANAI--VRLGSPSS--NS 393
 QY 295 VRG-----WAIDRADREFRLRLNMLFRGCPDPRRYLLQRFYRLPOPLIERF 343
 DB 394 LRQGLSAFWRDMPTRRRQRRFRCGMILK-LDLDATRRFDFAPFDLPQHYMGF 452
 QY 344 YAGRLTLAD-----RLRVTGRPPILSQAVRCIPERPLIOER 381
 DB 453 LSRRLFLBELVFGSLFSPHSANSTRLEIMT-KGVPLAKMI-----NNLVQDR 500
 RESULT 15
 CCS_CITSI
 ID CCS_CITSI STANDARD; PRT; 503 AA.
 AC Q9SEAO;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Capsanthin/capsorubin synthase, chloroplast precursor.
 GN Name=CCS;
 OS Citrus sinensis (Sweet orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eustros II; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=27111;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu C.J., Chen D.M., Zhang S.L.;
 RT "Molecular cloning of capsanthin/capsorubin synthase gene from orange
 (Citrus sinensis).";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Catalyzes the conversion of the ubiquitous 5,6-
 CC epoxycarotenoids, antheraxanthin and violaxanthin, into capsanthin
 CC and capsorubin, respectively.
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast; chromoplast (by similarity).
 CC -!- SIMILARITY: Belongs to the lycopene cyclase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 CC EMBL, AF169241; AAF18389.1; -;
 CC InterPro: IPR010108; Carotene_cycl.
 CC InterPro: IPR008671; Lycopene_cycl.
 CC Pfam: PF05834; Lycopene_cycl; 1.
 CC TIGFAMS: TIGR01790; carotene_cycl; 1.
 CC Carotenoid biosynthesis; Chloroplast; NAD; Oxidoreductase;
 CC Transist peptide.
 CC TRANSIT 1 2 Chloroplast (Potential).
 CC

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 19:18:35 ; Search time 33.0577 Seconds
(without alignments)
111.838 Million cell updates/sec

Title: US-10-695-980-5

Perfect score: 1999

Sequence: 1 MSHDLLIAGAGLGGALIALA.....IFLSQAVRCLPRLPLQGERA 382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	810.5	40.5	382	C37802	crty protein - Erw
2	795.5	39.8	382	S52585	lycopen cyclase -
3	750	37.5	386	S52581	lycopen cyclase -
4	162	8.1	410	D75475	lycopen cyclase -
5	161.5	8.1	500	S66350	lycopen beta-cycl
6	161.5	8.1	500	S72505	lycopen beta-cycl
7	161	8.1	500	S72506	lycopen beta-cycl
8	160	8.0	500	S66349	lycopen beta-cycl
9	155	7.8	405	T37022	probable lycopen
10	132	6.6	498	S71511	capsanthin/capsor
11	126.5	6.3	526	T07082	lycopen epsilon-c
12	125	6.3	408	D70549	hypothetical prote
13	122	6.1	402	AC2797	2-octaprenyl-6-met
14	122	6.1	402	C97576	2-octaprenyl-6-met
15	117	5.9	388	T36402	probable monooxyge
16	116	5.8	338	C70553	hypothetical prote
17	115.5	5.8	421	E97689	hypothetical prote
18	115.5	5.8	421	AB2915	FAD-dependent mono
19	114.5	5.7	937	D87483	ribonucleotide red
20	114	5.7	508	S44950	lmbc protein - Str
21	112.5	5.6	466	C83407	hypothetical prote
22	112.5	5.6	471	S51511	capsanthin/capsor
23	111.5	5.6	1021	H75423	hypothetical prote
24	111	5.6	391	G85566	hypothetical prote
25	111	5.6	406	T36632	probable oxidoredu
26	109.5	5.5	456	D84202	hypothetical prote
27	109	5.5	391	AD0584	probable monooxyge
28	109	5.5	412	A83461	hypothetical prote
29	108.5	5.4	1542	AC3173	conjugal transfer

30	108	5.4	391	2	D90716	oxygenase [impor
31	108	5.4	429	2	AH3272	2-octaprenyl-3-met
32	108	5.4	757	2	JC7726	(1->4)-alpha-D-glu
33	107	5.4	398	2	G82755	conserved hypothet
34	106	5.3	610	2	T06280	probable starch sy
35	105.5	5.3	408	2	H87193	probable FAD-lik
36	105.5	5.3	429	2	F84015	maltoase/maltodextr
37	105	5.3	337	2	D87354	conserved hypothet
38	105	5.3	391	2	D64801	probable monooxyge
39	104.5	5.2	476	2	T36395	probable transcript
40	104	5.2	475	2	T35799	probable oxidoredu
41	104	5.2	503	2	S55589	D-nopaline dehydro
42	104	5.2	549	2	JC7158	steroid monooxygen
43	103.5	5.2	392	2	AG0873	2-octaprenyl-6-met
44	103.5	5.2	472	2	AC0863	l-fuculose kinase
45	103	5.2	433	2	H87660	peptidoglycan-bind

ALIGNMENTS

RESULT 1

C37802
crty protein - Erw
C/Species: Erwina uredovora
C/Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 09-Jul-2004
C/Accession: C37802
R/Mitsawa, N.; Nakagawa, M.; Kobayashi, K.; Yamano, S.; Izawa, Y.; Nakamura, K.; Harashio, J. Bacteriol. 172, 6704-6712, 1990
A/Title: Elucidation of the Erwina uredovora carotenoid biosynthetic pathway by function
A/Reference number: A37802; PMID:91072214; PMID:2254247
A/Accession: C37802
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-382 <MIS>
C/Superfamily: Erwina dycopen cyclase

Query Match 40.5%; Score 810.5; DB 2; Length 382;

Best Local Similarity 45.6%; Pred. No. 3.5e-56;

Matches 171; Conservative 59; Mismatches 138; Indels 7; Gaps 4;

QY	3	HDLLIAGAGLGGALIALAVDRRPARIVMLDARGSPDQHTWSCHDTLSPEWTLARLSP	62
DB	5	YDLIIWAGLGLNLRLIQOOQPMKRLIIDAPQAGNHTWSFHDDLTSGHRWIP	64
QY	63	IRRGWTDGEVAFPHSRRLTTGYGIEAGLIGLQ--GVDLRNTHVATLDTGATL	119
DB	65	LVVHWPDYQVFPFRRLKNGYFCITSGRPAEVLQRFQPHLMWDIAVAEVNAESVRL	124
QY	120	TGGSRIEACVLDAGAVETPHLTVGFQFQVEVEIETDAPHGVPRMIMDATVPMQDGR	179
DB	125	KKGQVIGARAVIDGRVANSALSVGFQEWRLSHPHGISSTINDATVDDQNGR	184
QY	180	FYLLPFPSPTRLIEDTRYSDGDDLDGALAASLDYARRGMTQGE-WRRRGILPIAL	238
DB	185	FYSLPLSTRLIEDTHYIDNATIDPECARONICDYAAQGMQGLTLREGGALPIIL	244
QY	239	AHDAIGFWRDHAQAVPVGAGLPHVPTGSLPYAAQVADAIARDL-TTASARRAVNG	297
DB	245	SGNDAFWQGRPLAC--SGLRAGLFHPTTGYSLPLAVALDRLSALDVFTSAIHHAIRH	302
QY	298	MAIDRADRPRLPLNRLMFRGCPRRRLRLQRFRLPOPILIERVYARLTLAQLRLTV	357
DB	303	FRERWQOGFFRMRNRLFLAGPADSRWRVQRFYGLDELILARYAKLTTLTRLRL	362
QY	358	TGRPPPLSQAVRCL 372	
DB	363	SGKPPVPLAALQAI 377	
RESULT 2	SS2585		

QY 252 -GAVPVGLAGLPHPTGYSLPYAAQVAAIAARDLTTASAR---RAVGMAI-----DR 302
 Db 261 GGVLAVGAAGAAHVPVSGFQVAGALSDAGVATA-TATLCCGKDAAGAAALNSPERK 319
 QY 303 ADRREPLRLNMLFRCPPDRRYRLQRFYRLPQPLIERFY-----AGRLTLADRLRV 357
 Db 320 AARE--VHLGVGALLGRLRAELPHFPGFFGLPFRQWARFLPPTDAGTLA-RTMLRVF 376
 QY 358 T---GRPPIPLSQAVRCLP 373
 Db 377 AQTGRVRLPLARAAALAP 395

RESULT 5

S66350
 Lycopen beta-cyclase (EC 5.5.1.-) - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 28-Jul-2000
 C:Accession: S66350
 R:Pecker, I.; Gabbay, R.; Cunningham Jr., F.X.; Hirschberg, J.
 Plant Mol. Biol. 30, 807-819, 1996
 A>Title: Cloning and characterization of the cDNA for lycopen beta-cyclase from tomato
 A:Reference number: S66349; M01D:96194462; PMID:8624411
 A:Accession: S66350
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-500 <PEC>
 A:Cross-references: EMBL:X86452
 C:Superfamily: tomato lycopen beta-cyclase
 C:Keywords: intramolecular lyase; isomerase

Query Match 8.1%; Score 161.5; DB 2; Length 500;
 Best Local Similarity 21.7%; Pred. No. 5.1e-05;
 Matches 103; Conservative 70; Mismatches 146; Indels 155; Gaps 28;
 QY 4 DLLIAGAGLGAIA-----LAVRRRPARIVMLDARSGPSDQHTW-----S 46
 Db 85 DLAVVGGGAGLAVAAQVSEAGLSVCSIDPNFKLIW-----PNNYGVWVDEFEAMDLD 138
 QY 47 CHDTLSPEWMLRLSPFRGEWTDQVAFPDHSRR-LTTGYGSIKAGALIG-----LIQ 99
 Db 139 CLD-----ATWGAAYVIDDNTAKDLHRYGRVNRKQKSKMOKCINM 182
 QY 100 GVDLRKNTHTVATLDOTGATL---TDGSRTEACVIDARG-----AVETPHLTVGQKRV 150
 Db 183 GVKFHQAKVIKYIHESKSMILCNDGITITQATVVDATGFSRLVQYDPRY-NPGYQVAY 241
 QY 151 GV--ELETDPAPGVERPMID-----ATVPQMDGYRFTYLLPFPSTRIL 192
 Db 242 GILAEVE-EHPFDVNMKVMFMDRSHLKNNTDLKERNRIIP-----TFLYAMPFSSNRIF 295
 QY 193 IEDTRY--SDGGLDDGALAQAASLDYAARGMTQGMRR-----ERGILPILAHDAIGFW 246
 Db 296 LRETSIVARPGRLRID--IQERM--VARLNHLGIKYSIEDEHCLIPMG----- 341
 QY 247 RDHAQAVP-----VGLG--AGLPHPTGY---SLPYAAQVADAI-----AAR 284
 Db 342 -----GRLPVLPRQVVGIGTAGMVPSTGYVARTLAAPVANAIIQYLSERSHSGN 396
 QY 285 DLTTSARARAVGMWID-RADRDRLRLNMLFRCPPDRRYRLQRFYRLPQPLIERF 343
 Db 397 ELSTAVWKUL---WPIERRRQREFFCGMDILKLDLPATRRF--FDAFDELEPRYWHGF 451
 QY 344 YAGRLTLAD-----RLRIVTGRPPIPLSQAVRCLPERPILQER 381
 Db 452 LSSRLFLPELIVFGSLFSHASNTSRFEIMT-KGTVPVNMVMI-----NNLLQDK 499

RESULT 6
 S72505
 Lycopen beta-cyclase (EC 5.5.1.-) - tomato
 C:Species: Lycopersicon esculentum (tomato)

C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 09-Jul-2004
 C:Accession: S72505
 R:Pecker, I.; Gabbay, R.; Cunningham Jr., F.X.; Hirschberg, J.
 submitted to the EMBL Data Library, April 1995
 A>Description: Cloning and characterization of the lycopen cyclase gene from tomato.
 A:Reference number: S72505
 A:Accession: S72505
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-500 <PEC>
 A:Cross-references: UNIPROT:Q43503; EMBL:X86452; NID:g1006672; PID:g1006673
 C:Superfamily: tomato lycopen beta-cyclase
 C:Keywords: intramolecular lyase; isomerase

Query Match 8.1%; Score 161.5; DB 2; Length 500;
 Best Local Similarity 21.7%; Pred. No. 5.1e-05;
 Matches 103; Conservative 70; Mismatches 146; Indels 155; Gaps 28;
 QY 4 DLLIAGAGLGAIA-----LAVRRRPARIVMLDARSGPSDQHTW-----S 46
 Db 85 DLAVVGGGAGLAVAAQVSEAGLSVCSIDPNFKLIW-----PNNYGVWVDEFEAMDLD 138
 QY 47 CHDTLSPEWMLRLSPFRGEWTDQVAFPDHSRR-LTTGYGSIKAGALIG-----LIQ 99
 Db 139 CLD-----ATWGAAYVIDDNTAKDLHRYGRVNRKQKSKMOKCINM 182
 QY 100 GVDLRKNTHTVATLDOTGATL---TDGSRTEACVIDARG-----AVETPHLTVGQKRV 150
 Db 183 GVKFHQAKVIKYIHESKSMILCNDGITITQATVVDATGFSRLVQYDPRY-NPGYQVAY 241
 QY 151 GV--ELETDPAPGVERPMID-----ATVPQMDGYRFTYLLPFPSTRIL 192
 Db 242 GILAEVE-EHPFDVNMKVMFMDRSHLKNNTDLKERNRIIP-----TFLYAMPFSSNRIF 295
 QY 193 IEDTRY--SDGGLDDGALAQAASLDYAARGMTQGMRR-----ERGILPILAHDAIGFW 246
 Db 296 LRETSIVARPGRLRID--IQERM--VARLNHLGIKYSIEDEHCLIPMG----- 341
 QY 247 RDHAQAVP-----VGLG--AGLPHPTGY---SLPYAAQVADAI-----AAR 284
 Db 342 -----GRLPVLPRQVVGIGTAGMVPSTGYVARTLAAPVANAIIQYLSERSHSGN 396
 QY 285 DLTTSARARAVGMWID-RADRDRLRLNMLFRCPPDRRYRLQRFYRLPQPLIERF 343
 Db 397 ELSTAVWKUL---WPIERRRQREFFCGMDILKLDLPATRRF--FDAFDELEPRYWHGF 451
 QY 344 YAGRLTLAD-----RLRIVTGRPPIPLSQAVRCLPERPILQER 381
 Db 452 LSSRLFLPELIVFGSLFSHASNTSRFEIMT-KGTVPVNMVMI-----NNLLQDK 499

RESULT 7

S72506
 Lycopen beta-cyclase (EC 5.5.1.-) - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 09-Jul-2004
 C:Accession: S72506
 R:Pecker, I.; Gabbay, R.; Cunningham Jr., F.X.; Hirschberg, J.
 submitted to the EMBL Data Library, September 1994
 A>Description: Cloning and characterization of the lycopen cyclase gene from tomato.
 A:Reference number: S72506
 A:Accession: S72506
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-500 <PEC>
 A:Cross-references: UNIPROT:Q43578; EMBL:X81787; NID:g1006689; PID:g1006690
 C:Superfamily: tomato lycopen beta-cyclase
 C:Keywords: intramolecular lyase; isomerase

Query Match 8.1%; Score 161; DB 2; Length 500;
 Best Local Similarity 22.7%; Pred. No. 5.6e-05;
 Matches 105; Conservative 62; Mismatches 151; Indels 144; Gaps 27;

```

OY      4  LLLLAGAGLSGGLILAAVRDRPRDARIWMLDRSGSPDOHTSGCHDTLSPFWMLARLSPI 63
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      85  DLAVVGGGPGAGLAAVOVSE--AGLSVSIID--PSPKLIW-----PN 122
OY      64  RRGWTDQ-----EVAPPDH--SRLLTGYGSIAGALLG-----LIQ 99
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      123  NYGVWVDFEFAEMDILDCLATWSGTVVYIDBWTXKDLBRPYGRVNRKQJCKSKMOKCILI 182
OY      100  GVDLRKMNTHVATLDTGTATL--TDGSRLEACVUIDAG-----AVETPHLTGVPQKPV 150
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      183  GVKPFHAKVIKIVIEEAKSMLICNDGVTIQATVVLADATGFSKCLVQYDQPY-KPQYQAVY 241
OY      151  GV-:EITDPAHPGVPRMIMD-----ATVQMDGGRFYIYLLPFSPTRI 192
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      242  GLIAEVE-EHPDTSKMYLMDRDSHLGNMMLKERNKRV-----TPIYAPFSSNKIF 295
OY      193  IEDTRY--SDGGDLDGALAQAQLDYAARGTGOEMRR---ERGILPITALHAIGFW 246
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      296  LEETSLVAPRGIRMD--IQERM--VALNLHGIKVKSIEEDECVIIPMG----- 341
OY      247  RDHAGAMP-----VGLG--AGLFHPVGY--SLPFAQVQAAI-----AARDLTTA 289
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      342  ----GSLPVIQORVVGSTAGLVHPSTGYVAVARTLAAAPVANAIIHYLQSEKDLGN 396
OY      290  SARRAVRG--WAID-RADRDREPLRLNLNMLFRGCPDPRRYLLQRFYRLPQPLIERFYAG 346
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      397  ELAAVWMDLWPIERRQREFFCFGMDILKLKDLPATRRF--FDAFPELDERYHWGFLSS 454
OY      347  RLTLAD-----RLAIIVGRPIPIPSQAVRCL 372
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      455  RLVIPELIFPGSLFSRASNTSRIEIMT-KGTLPLVNNMINNL 495

```

```

Db      296 LEETSLVAPGCLRLDD---IQERM--VARLNHGIIKKSIEEDENHCYIPMG----- 341
QY      247 RDHQGVNP-----VGLG--AGLEHPTGY----SLPYAAQVADAI-----AARDLTVA 289
Db      342 ----GSLPEIPQREVGTGTAGLVHPSTGIMVARTAAAPVANAIIHLYSEKDLGN 336
QY      290 SARBAVNG--WAID-RADRDREFLNLNMLFRGCPDRRYRLLOREYRLPOPLIERPYAG 346
Db      397 ELSAVWMDLPIPIRRRQRQREFFCGMDILKLDLPAFRRE--PDAFDLEPRYWHGFLSS 454
QY      347 RLTLAD-----RLRIYGRPRPIPLGAVNCL 372
Db      455 RLYPELIFGLSLFSPASNTSRIEIMT-KGTLPVNMIMNL 495

RESULT 9
T37022
probable lysozyme cyclase - Streptomyces coelicolor
C.Species: Streptomyces coelicolor
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C.Accession: T37022
R.Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A.Reference number: Z21619
A.Accession: T37022
A.Status: preliminary; translated from GB/EMBL/DDBJ
A.Molecule type: DNA
A.Residues: 1-405 <MR>
A.Cross-References: UNIPROT:Q9RI55, EMBL:AL109989, PDB:CAM53415.1, GSPDB:GN00070; SCOPED
A.Experimental source: Strain A3(2)
C.Genetics:
A.Gene: crty; SCOPDB:SCU12.03c

```

```

RESULT 8
S66349
lycopenene beta-cyclase (EC 5.5.1.-) - common tobacco C;Species: Nicotiana tabacum (common tobacco) C;Date: 14-Feb-1997 #sequence revision 13-Mar-1997 #text_change 28-Jul-2000 C;Accession: S66349 Ripecker, I.; Gabbay, R.; Cunningham Jr., F.X.; Hirschberg, J. Plant Mol. Biol. 30, 807-819, 1996 A>Title: Cloning and characterization of the cDNA for lycopenene beta-cyclase from tomato A.Reference number: S66349; MUID:9619446z; PMID:8624411 A;Accession: S66349 A>Status: preliminary; nucleic acid sequence not shown A:Molecule type: DNA A:Residues: 1-500 <PEC> A:Cross-references: EMBL:X81787 C:Superfamily: tomato lycopenene beta-cyclase C:Keywords: intramolecular lyase; isomerase

Query Match          8.0%; Score 160; DB 2; Length 500;
Best Local Similarity 22.5%; Pred. No. 6,7e+05;
Matches 104; Conservative 62; Mismatches 152; Indels 144; Gaps 26;

Qy      4 DILLGAGISGLIALAVRRPRDARIWLDARSGPSDDHTSCHDTLSPEWLRLSTI 63
         |||::||::||::||::||::||::||::||::||::||::||::||::||
Db     85 DLVAVGGGPGAGIAVAQOVSE--AGLSVSID-----PSPKLIV-----FN 122

Qy    64 RRGEWTDEVFAP-----PDHSRRLTTGVGSIEAGALIG---LLQ   99
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   123 NYGVAVNVEEPAADLLDCDLATWSGVVVYIIDNTTKDLDPYGRVNKRKQLKSROMOKCIIM 182

Qy   100 GVVDLRMNTHVALTLDPTGTALT--TDGSRLEACVIDARG-----AETPHLTVGFOKEV 150
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   183 GVKFHAAKIYIKIHBEAKSMILCNDGVTIQAITVLDAITFSCFLVOYDYEPY-KPGYQNAV 241

Qy   151 GV-EIEDPARHVERPMIMD-----ATVPOMDGRYFTYLLPFSPTRLI 192
        ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   242 GLTAEEV-EHPEDTSRMVLMDWRDSLGNMELKERNRKV----TFLYAMPFGSNRIKF 295

Qy   193 IEDTRY--SDGDSDLDAQSLSYAARGMTGCAMRA---ERCILFIALAHAIGAIFW 246
       :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

Query Match          7.8%; Score 155; DB 2; Length 405;
Best Local Similarity 22.9%; Pred. No. 0.00013;
Matches 105; Conservative 45; Mismatches 155; Indels 154; Gaps 23;

QY      DLLIAGLGLSALTAIAVDRRPPARIIVMLDARSGP--SDQHWSCHDTDLSPMTABLS 61
       |||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      DVLVIIGGAALSLAHLLTENGTAPMATILVEPRPGPLRPARRTW-CY----- 55
QY      62 PIRGEWTDGEVAPDPHSHRLTTGYGSIIEGALIGLVDRMNT--HVA-----TLD 113
       |               |||::||::||::||::||::||::||::||::||:
DB      56 -----W-----GAADGLEAVGASWSVLRLLHGADGSVTVD 87
QY      114 DTGATL-----TDSRI-----EAC-----129
DB      88 PAPFYRWVSADPERHWGHRLARTDGARLLRGTAESRAVPAGTEVRCTLPGRPTLY 147
QY      130 ---YTDAKGATEPHL-TVGFQKEVG--VELETDA--PHGVSRPMIDATVPO-MDGYSF 180
       |||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      148 ARRVFDSPLDELPPARCCLIQHRGMVFVTRTRDPRAVAD--LMDPVPOPAHGLAF 204
QY      181 IYLFPFSPTRIIEDTVSDG---GDLDGGLAQASILDYAARKGMTQEQEMRRERGLPIA 237
       |||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      205 GYVFLPADRLVLTETESSRAPLTTEAVEGSLGHYCDDIIIGLGELTYE--RTQGVIPTM 262
QY      238 LAHAIDIGFWHDQAQAV-PVGLGAGLFHPVTGYSILPYAAOVADAIAA--RD----LTTAS 290
       :               |||::||::||::||::||::||::||::||::||:
DB      263 DAR-----FPCRAPAVAYRIGTAGCATRPATGTGYFAVAQHSNGALIAALLDHGDHVRVAPH 317
QY      291 ARRA-----VRGMAIDRADRDREFLLRNMLEFRGCPDRRRRYLLORFRLRPOPILIERF 343
       |||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      318 GRRARAMDVAVLRALDTRIDGRIDGPRFFPTDL-----FRVPAERILLRF 358
QY      344 YAGRUTTLADRILYITGRPIPLSQAVRCLEPRLIOERA 382
       |||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      359 LDGTTSLREEWGIGLTPVRPM--LRTAAVPPLPFRSS 394

RESULT 10
871511
capsanthin/capsorubin synthase (EC 5.5.1.-) - pepper

```

N:Alternate names: chromoplast-specific protein
N:Contains: capsorubin synthase; capsanthin synthase
C:Species: Capsicum annuum (pepper)
C>Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 09-Jul-2004
C/Accession: S71511; J02141
R:Bouvier, F.; Huguency, P.; d'Harlingue, A.; Kuntz, M.; Camara, B.
Plant J. 6, 45-54, 1994
A>Title: Xanthophyll biosynthesis in chromoplasts: isolation and molecular cloning of an
A/Reference number: S71511; MUID:9504653; PMID:7920703
A/Accession: S71511
A:Molecule type: mRNA
A/Residues: 1-498 <BOU>
A/Cross-references: UNIPROT:Q42435; EMBL:X76165; NID:9522119; PIDN:CAAS3759.1; PID:9522212
A/Experimental source: cv. Lamyo; developmental stage ripening fruits
R:Denervet, U.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.
Biochem. Biophys. Res. Commun. 199, 1144-1150, 1994
A>Title: Structure and expression of two plant genes encoding chromoplast-specific protease
A/Reference number: JC2140; MUID:94197697; PMID:8147854
A/Accession: JC2141
A:Molecule type: DNA
A/Residues: 1-498 <DEB>
A/Cross-references: EMBL:X77289; NID:9468747; PIDN:CAAS4495.1; PID:9468748
A/Experimental source: cv. Yolo wonder
R:Denervet, U.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.
Biochem. Biophys. Res. Commun. 201, 466, 1994
A/Reference number: A55775
A/Contents: annotation; erratum
A>Note: The legends of the nucleotide sequence figures were reversed in publication (see
C:Genetics:
A:Gene: CCS
C:Complex: monomer
C:Function:
A>Description: catalyses the conversion of the ubiquitous 5,6-epoxycarotenoids, anthraxax
A:Pathway: carotenoid biosynthesis
C:Superfamily: tomato lycopene beta-cyclase
C:Keywords: chromoplast; intramolecular lyase; isomerase

Query Match 6.6% Score 132; DB 2; Length 498;
Best Local Similarity 19.6%; Pred. No. 0.011;
Matches 89; Conservative 68; Mismatches 154; Indels 144; Gaps 20;

Db 4 DLIAGAGLSGALIALAVDRRPARIVMLDARSGSPSDHTWSCHDTDLSPEWLARLSP I 63
 :::|||:::||||:|:::|:::|:::|:::|:::|:::|:::|
83 DVIIIGGPAG--LRLAQVSKYGIKVCVD-----PFLSLMW-----PN 120

Db 64 RRGEXTOE-----VAFPDH-SRRLLTGGISIEAGAL-IGLIQG-VDL 103

Db 121 NYGVWVDEFEEKLGEDCLDHKWPVSCVAISHDKTKYLDRPYGRVSRKKLKILNSCVEN 180

Oy 104 RMNHVATL-----DDTGATITDGSRIEAACVIDANG-----AVETPLTVTGFOKEV 150

Db 181 RVKFYKAVLVKYGHFEFFSSIVCDGKRKISGLVLVDASGYASDIYEVDKP-NHGVOVAH 239

Oy 151 GVRIETD-APGVERPMID-----ATVPOMDGYRFIYLPSPSPRLIEDTRY 198
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
240 GILAEVDNDHPDDLDMKMLMDWRSHLGNEPLYLRVKNTEPPFLYAMPDRNLVLEEF-- 297

Db 199 SDGDLDLGAQAQSUDY-----AARGMTGDEM---REGGILPIALAHAIQEMR 247
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
298 -----SLVSREPLSTMEVKRMVVALRHGIKRVSLVEEKCVTIMGSPRLRI---- 345

Oy 248 DHAGAVPVGAGAGLFHPVTGSLPYAAQVDAIAARLTATTAARRAVRWAIADRADRD 307

Db 346 --PQNVAIGGTSGIVHPSGSCTMAKSMAALPVLAALVESLGSTMIRG----- 393

Oy 308 FLRLINRLFGCCPEDDR-----YLLORFYRLPOPLIERFYG 346
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
394 -SOLVHRVMNGIMPEDRRVRRECFCFMETILLKLDELSTRRLPALFAFDVDEPKWMGFSS 452

Oy 347 RLTLAD-----RLRTVGRRPIPI 365
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
453 RLSVRELAVLSLYLFGHASNARDLDTV-KCTVPL 466

RESULT 11

107082

Lycopene epsilon-cyclase (EC 5.5.1.-) - tomato

C1Species: Lycopersicon esculentum (tomato)

C1Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004

C1Accession: T07082

R1Name: G.J. Cohen, M.J. Zamir, D.J. Hirschberg, J. submitted to the EMBL Data Library, July 1997

A1Description: Regulation of expression of the gene for lycopene epsilon cyclase during

A1Accession: T07082

A1Status: preliminary; translated from GB/EMBL/DBD

A1Molecule type: mRNA

A1Residues: 1526 <RON>

A1Cross-references: UNIPROT:O65837, EMBL:Y14387, NID:e1285211, PIDN:CAA74745.1, PID:e128

A1Experimental source: cultivar VF36; leaf

C1Genetics:

A1Gene: CrtI-e-1

A1Map position: 12

C1Function:

A1Description: converts lycopene to delta-carotene

C1Superfamily: tomato lycopene beta-cyclase

C1Keywords: intramolecular lyase; isomerase

Query Match 6.3%; Score 126.5; DB 2; Length 526;

Best Local Similarity 21.1%; Pred. No. 0.031;

Matches 89; Conservative 55; Mismatches 164; Indels 113; Gaps 20;

4 DLLIAGAGLSCALIALAVDRRPPARIIVMLDARSGP-----SDQHTWSCDPTDISPEMLA 58

107 DLVVITGGCGPAG--LALAASAKGLANGLV-----GPDLPFTNNGVMEDEPKDGLG----- 155

59 RLSTIRGGEWTDQVAPRPHSRRLTTG--YG-----STEAGALIGLQGV 102

156 -LQACIEHWMD-TIVYLDDEPTLIGRAYGRVSRHFLHBEILKRCVEAGVL----- 205

103 LRMTHTVATLDDTGTATLTDGSR1---EAACYTDAR-----GAVETPHL-----TVGF 146

206 -----YINSKYKDRIVEATNGSLVECGDVIIPIRPFVTASGASGKFLQYELGSPRVSV 260

147 QKFWGVEIETD-APHGVERPMIDATVTPQMDGYR-----FYTLPPSPPT 189

261 QTAVGVEVEVDNPNFDSLWVMD-----YRDYLRHQAQSLKAYPTFLYMPMSPT 312

190 RILIEDPRYSUGGDLDDGALAQAASLDYARRGWTCQEMRER-GILPITALHDLGFWRD 248

313 RVFEETCTLASKDMPPDLKKKLMRLNTLGVRIKEIYEBMSYIPVG-----GSLPN 366

249 HAQGAVPVGAAGLGFHPVTGY---SLPYAAQVADA1-----ARDLTASARRAVRG 297

367 TEQKTLAFGAAMSVHRTGYSVRSLSERPKCASVLANILRQHSQMLTSSISPIST 426

298 -----WAIDRADRDLRLNRLMFLRCGPPDRRYLLQRRYRLPQPLIRFPYAGRLTLA 351

427 QAMNTLWPOER-KKQRSFFLFGALILQLDIEGRISFFRAFFRVPMWMOGFLGSSLSA 485

352 D 352

486 D 486

RESULT 12

D70549

hypochemical protein Rv0561c - Mycobacterium tuberculosis (strain H37RV)

C1Species: Mycobacterium tuberculosis

C1Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C1Accession: D70549

R1Name: S.T. Broesch, R.J. Parkhill, J.J. Garnier, T.J. Churche, C.J. Harris, D.J. Gordon, S.J. Connor, R.J. Davies, R.J. Davlin, K.J. Feltwell, T.J. Gentles, S.J. Hamlin, N.J. Holroyd, S. Rajandream, M.A.I. Rogers, J.J. Rutter, S.J. Seeger, K.J. Skelton, S.J. Squires, S. Nature 393, 537-544, 1998

A1Authors: Squires, R.J. Skelton, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70549
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-408 <COL>
A;Cross-references: UNIPROT:006427; GB:Z95558; GB:AL123456; NID:g3261781; PIDN:CAB08972
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0561c
A;Superfamily: fixc protein

```

Query Match      6.3%; Score 125; DB 2; Length 408;
Best Local Similarity 21.3%; Pred. No. 0.02%;
Matches 99; Conservative 56; Mismatches 142; Indels 168; Gaps 23;

QY      2 SHDLIIAAGISGLIALAVRDRPRDARIWMLDARSGPSDDHTWSCHDTDISPEWLAITS 61
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db      6 SADVVVVGAGPRGS--AAAAMAAARAGRDVLVIDTATFPFDK---PCGD-GILPRAVAELH 59
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
QY     62 PIRGSEWTDQEAVPFDHSR---RLTTGYSIEA-----GALIGLLQGVDIRMNT 107
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db     60 QLGCGKW-----LADHIRHRGRLMSGCGEVEVDWPPSPSYCSAAVRLK-LDDRIR- 111
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
QY    108 HVAATLDDTGA-----TLTDSRIEAAVCYIDARGAETP----- 140
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db    112 --KVAEDTGAMLLGAKAAVAVHHDSRRVSLTTLADGTEVGCGRLIVADGA-RSLPGRKL 168
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
QY    141 -----HLTVGF--QKPYGVELEIDAPRGVERPMTMDATVQMDGRTY 182
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db    169 GRMWRRETIVGVAAVRGYSTAYSDPWLTHSHELRSPPGAVLP-----GYGWIIF 217
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
QY    183 LLPSPTIILEDVRYSDGDLDDGAL--AQAALDYAARR-----GWTQGEW 227
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db    218 PLNGEENVIGV-----GALSTSRKPPADLRPLLIYSTTDLRRDEWGTGQPR 264
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
QY    228 RRERGIPLIALAHDAIGFMRDHAQGAVPVGLGAGLFHEVTVGSYLEPAAQ---VADAIAA 283
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db    265 AVSSALLPLFMGAVSGVA-----GSNWMLIGDAAACVNPFLNGEGIDYGLJETRIAAELIDS 319
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
QY    284 RDLTTASARARVRCMAIDRADDRFLRLINRLFRGCPDDRRIYLLQRFYR-LP----- 336
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db    320 RDL-----ARLWPSLADR-----YGRGFSVAARLALLLTFPFRFLPTTGPIT 361
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
QY    337 -----QPLIERFYAGRLTLADR-----LAIYVRGP 362
      | : : : | : : | : : | : : : : : : | : : | : : | : : |
Db    362 MRSTALMNIAVRVMSNLVTDDRDWRVAARVWVGCGQLSLTVRRRP 406

```

RESULT 13
AC2797
2-octaprenyl-6-methoxyphenol hydroxylase [imported] - *Agrobacterium tumefaciens* (strain
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AC2797
R/Mood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, P.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AC2797
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 17402 <KUP>
A/Cross-references: UNIPROT:O80UG3; GB:AE008668; PIND:AA42793.1; PID:g17740238; GSPDB:0
C/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: ubiH
A/Map position: circular chromosome

Query Match	6.1%;	Score 122;	DB 2;	Length 402;
Best Local Similarity	22.0%;	Pred. No. 0.05;		
Matches	98;	Conservative	54;	Mismatches 166;
			Indels	128;
			Gaps	20;

```

QY      1 MSH-DLLIAAGCSGLIALIAAVDRRPAIIVLIDARSGSDQHTMSCHD----- 49
Db      1 MRPHDIAITAGAGQIAAIAL--ARAGRHVALLIDSEKKDRTTALMDQSIIRFMDRLG 58
QY      50 --TDLSEPMIARLSPIRGEWTQDEVAFPDHSRLTTGYGSIAG-----ALIG 96
Db      59 LMSIARS-AARSTWQIIDGTDRLRAP-----TVQFRSSSEIGDAFGWNIENEAULT 111
QY      97 LL-OGVDLRNMTHTVAITDDT-----GATLIDGSRIEAACVINDAGAVETHLTIV 144
Db      112 VLSEAVEQEHN--ITRLDITTAETIIDIGNRISVTIADGEVLSADFLIGADGRKSMVREA 169
QY      145 GF-----QKPFVEIETDAPHG-----VERPMIDATVPQMDOG 177
Db      170 GIGYKMSYQYQYAVLVNFHSRPHGNVSTFPHPTGFTQVPLPDRBSLSLVVYVTEQQ-- 227
QY      178 YRFYLLPFSPTRIILIEDTRYSDGGDLDDGALAQASLIDVAARRGWTQGOENRRER-GILPI 236
Db      228 AAEITALPLETILSLKVEERMQSMLGAVTVENSVOA-----WPLSSWTAHFGKGRV 278
QY      237 ALAMDALGFMRDHAQGAVPVGLAGLPHPTGYSLPIAAQVADAIARDDTTASARRAVR 296
Db      279 ALIIEA-----AHGPPPIG-----AOGNLT-----SLRDIATELLGAVSROPPIA 319
QY      297 GMAIDRADRDRLFLLIRMLFRGCPDRRYRLIQRFYRLPOPLIERPYAGRLTLADRLRI 356
Db      320 ADASSSTDRKRRAADVRSGLT-----SVDDLNR-----SLSDPLVPQMARAAAGLHV 365
QY      357 VTG-----RPPIPLSOAVRCLP 373
Db      366 LSGIGTLRSVMWRGIEPRGKLXALP 391

```

```

RESULT 14
C97576
2-octaprenyl]-6-methoxyphenol hydroxylase, ubiH (AF155772) [imported] - Agrobacterium tum
C|Species: Agrobacterium tumefaciens
C|Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C|Accession: C97576
R|Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ounullo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappae, C.; Markelz, B.;
Science 294, 2223-2328, 2001
A|Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A|Reference number: A97359; MUID:21608551; PMID:11743194
A|Accession: C97576
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-402 <KUR>
A|Cross-references: UNIPROT:Q8UEG3; GB:AE007869; PIDN:AAK87564.1; PID:G15156902; GSPDB:C
A|Gene: AGR_C_3301
A|Genetics:
A|Map position: circular chromosome

Query Match 6.1%; Score 122; DB 2; Length 402;
Best Local Similarity 22.0%; Pred. No. 0.05;
Matches 98; Conservative 54; Mismatches 166; Indels 128; Gaps 20;

OY 1 MSH-DLIAAGAGSGLALIAVDRDPRAIIVMLDARGSDQHTWSCHD----- 49
DB 1 MRHFDIATITGAGLAIATAIAL--ARAGRVAIALIAPSEKKDDRTALMDQSIREMDRLG 58
OY 50 --TDSPEMLARLPIRGSEWTDQEVAFPHDSRLTLFGYSIRAG-----ALIG 96
DB 59 LMSLIAAS-AAKRLSTMQIIDOTRLRAP-----IVQFSSRLGIDAFPGMNIPNEALIG 111
OY 97 LL-QGVDLRMNTHVATLDDT-----GATLTDGSIIEAACYADARGAVETPHLTIV 144
DB 112 VLSAIVQCEHN--ITRIIDTTAETIDICNDRIISTVIAIDGVLTSADFLIGADGRKSMVREAA 169

```

Job time : 36.0577 secs

```
QY 145 GF-----QKFGVEIETDAPHG-----VERPMINDATVPQMDG 177
DB 170 GIGVSWSYPTQAVTLNPSHSRPHGVSTEFHTPTGPTQVPLPDDRSSLVWVTPQ-- 227
QY 178 YRFIYLLPSPTRILLIEDTRYSDGDDLDGALAQAASLDYAARRGWTGQEMRRER-GILPI 236
DB 228 ABEITLPLPETLSLKVEEWMQSMIGAVTENSVOA-----WPLSMTAHRFGKGRV 278
QY 237 ALAHDAIGFWRDHAQAVVGLGAGLFHVTGYSLPYAQVADATIAARDLTASARRAVR 296
DB 279 ALIGEA-----AHGPPPIG-----AQGLNL-----SLRDITALTTELGAVSDFPIA 319
QY 297 GWAIDRADRDRLRLNRLFRGCCPPRRRYRLQRFYRLPQPLIERFYAGRLTLADRLRI 356
DB 320 ADAGSSFDKRRADYVSRLL-----SVDLINR-----SLISDFLPVQMARAGLHV 365
QY 357 VTG-----RPPILSQAVRCLP 373
DB 366 LSGIGTLRSMVMREGIEPGRGLKALP 391
```

RESULT 15

```
T36402
probable monoxygenase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36402
R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21606
A:Accession: T36402
A>Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-368 <SAU>
A:Cross-references: UNIPROT:Q9RK22; EMBL:AL109974; PIDN:CAB53314.1; GSPDB:GN00070; SCOE
A:Experimental source: strain A3(2)
C:Genetic8:
A:Gene: SCOE:SCF34.03c
```

Query Match 5.9%; Score 117; DB 2; Length 388;

Best Local Similarity 23.4%; Pred. No. 0.12;

Matches 92; Conservative 52; Mismatches 132; Indels 118; Gaps 22;

```
QY 7 TAGAGLSGL-----IALAVRRRPARIVMLDARSGPSDOH---TWSCHDTDS 53
DB 13 IIGAGGGITLCARVLRHGLAVTVYDRDPDA--ASRDQGSIDLHEEDGQLAREAGLI 69
QY 54 PEWLARLSPIRGEWTDQEVAFPPDHSRL-----TTGYGSIKAGALIG--LLQGV-- 102
DB 70 EEPFA-----LARGE-SQERRRFTAGRLGRLLPDEGETARPRIDRGQLRGILLIESLDAG 124
QY 103 -LRWNTHTVATLDDTG-----ATLTDSRIEACVIDARGAV-----TPHLT-VGF 146
DB 125 TVRWGHLGSVSGPAPGPRTLFTDGSIVETDLVIGDGAFSRVRAAVSDAVPRYTGVSF 184
QY 147 QKFGVEIETDAPHGVPRPMINDATVPQMDGYRFIY-----LPPSPTRILLIEDTRY 198
DB 185 LEAMFDMESAHPELSE--LVGRSAHVADGGQGLPAQRNSGGMRYVWRRVAL--DMWT 241
QY 199 SDG--GDLDDGALAQASLDYAAARGWTGQEMRR-----ERGLIPTALAHDAIGFWR 247
DB 242 ASGLRPDDTDGIRARLLAEYA--GMSPRILRMITENDGYVDRPLPALPVPHY-----WR 294
QY 248 -----DHAQAVPVGIGAGLFHPTGYSLPYAQVADATIAARDLTASARRAVR-- 296
DB 295 PTFGVTLTGDAHLMPLGIVNL-----AMLDGAEIALALAA-----SATVDDAVRTY 343
QY 297 -----GWAIDRADRDRLRL 311
DB 344 EKTMLPRSAEIAQMLEGAGFLLLEPDAEDLRL 377
```

Search completed: June 17, 2005, 19:38:04

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 19:37:17 ; Search time 130.394 Seconds
(without alignments)
1124.898 Million cell updates/sec

Title: US-10-695-980-5

Perfect score: 1999

Sequence: 1 MSHDLIAGAGLGGALIALA.....IFLSQAVRCLPERPLQERA 382

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	382	9	US-09-547-267-7
2	1999	100.0	382	10	US-09-920-923-5
3	1999	100.0	382	10	US-10-695-980-5
4	783.5	39.2	382	15	US-09-941-947A-10
5	783.5	39.2	382	14	US-10-218-118-6
6	783.5	39.2	382	15	US-10-363-567-30
7	783.5	39.2	382	15	US-10-466-656-4
8	783.5	39.2	382	16	US-10-735-019-6
9	783.5	39.2	382	16	US-10-734-778-6
10	783.5	39.2	382	16	US-10-735-442-6
11	783.5	39.2	382	17	US-10-848-307-4

12	783.5	39.2	382	17	US-10-886-906-6	Sequence 6, Appli
13	766	38.3	389	16	US-10-808-807-6	Sequence 6, Appli
14	760	38.0	387	17	US-10-810-723-8	Sequence 8, Appli
15	759	38.0	388	16	US-10-804-677-6	Sequence 6, Appli
16	757	37.9	388	16	US-10-808-979-6	Sequence 6, Appli
17	195.5	9.8	420	16	US-10-451-793-59	Sequence 59, Appli
18	195.5	9.8	501	9	US-09-323-998D-55	Sequence 55, Appli
19	195.5	9.8	515	15	US-10-425-114-40728	Sequence 68, Appli
20	191.5	9.6	419	16	US-10-451-793-68	Sequence 69, Appli
21	177	8.9	418	16	US-10-451-793-69	Sequence 8560, Ap
22	174.5	8.7	397	14	US-10-156-761-8560	Sequence 61, Appli
23	170.5	8.5	418	16	US-10-451-793-61	Sequence 57, Appli
24	170.5	8.5	498	9	US-09-323-998D-57	Sequence 63, Appli
25	163.5	8.2	502	16	US-10-451-793-63	Sequence 56, Appli
26	163.5	8.2	502	9	US-09-323-998D-56	Sequence 60, Appli
27	161.5	8.1	418	16	US-10-451-793-60	Sequence 58, Appli
28	161.5	8.1	500	9	US-09-323-998D-58	Sequence 62, Appli
29	161	8.1	418	16	US-10-451-793-62	Sequence 59, Appli
30	161	8.1	500	9	US-09-323-998D-59	Sequence 66, Appli
31	159	8.0	418	16	US-10-451-793-66	Sequence 64, Appli
32	157.5	7.9	422	16	US-10-451-793-64	Sequence 60, Appli
33	157.5	7.9	511	9	US-09-323-998D-60	Sequence 237207,
34	156.5	7.8	370	16	US-10-425-115-237207	Sequence 20, Appli
35	150.5	7.5	376	14	US-10-128-713A-20	Sequence 50, Appli
36	150.5	7.5	376	17	US-10-886-906-50	Sequence 67, Appli
37	145	7.3	418	16	US-10-451-793-67	Sequence 8, Appli
38	139	7.0	451	17	US-10-866-089-8	Sequence 65, Appli
39	137	6.9	417	16	US-10-451-793-65	Sequence 61, Appli
40	137	6.9	503	9	US-09-323-998D-61	Sequence 276683,
41	136	6.8	493	15	US-10-424-559-272683	Sequence 52568, A
42	131.5	6.6	353	15	US-10-425-114-52568	Sequence 20, Appli
43	127	6.4	456	9	US-09-323-998D-20	Sequence 75, Appli
44	126.5	6.3	422	16	US-10-451-793-75	Sequence 53, Appli
45	126.5	6.3	526	9	US-09-323-998D-53	

ALIGNMENTS

RESULT 1
US-09-547-267-7
Sequence 7, Appli Application US/09547267
Patent No. US2002014737A1
GENERAL INFORMATION:
APPLICANT: Hohmann, Hans-Peter
APPLICANT: Passamontes, Luis
APPLICANT: Tessier, Michel
APPLICANT: van Loon, Adolphus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/547,267
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: RAN 6002/170

TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-547-267-7

Query Match 100.0%; Score 1999; DB 9; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.1e-193;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLLIAGAGISGALLIAVDRRPPDARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARL 60
DB 1 MSHDLLIAGAGISGALLIAVDRRPPDARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARL 60
QY 61 SPIRGEMTDEVAFPDHSRRLTTGYSIEAGALIGLQGVDLRNMTHVATLDDTGATLT 120
DB 61 SPIRGEMTDEVAFPDHSRRLTTGYSIEAGALIGLQGVDLRNMTHVATLDDTGATLT 120
QY 121 DGSRIEACVIDARGAVETPHLTVGFQKFGVEIETDAPHGERPMINDATVPQMDGYRF 180
DB 121 DGSRIEACVIDARGAVETPHLTVGFQKFGVEIETDAPHGERPMINDATVPQMDGYRF 180
QY 181 IYLLPFSPTRLIEBTRYSDDGDLDDGALAQASLDYARRGWTGQEMREREGILPIALAH 240
DB 181 IYLLPFSPTRLIEBTRYSDDGDLDDGALAQASLDYARRGWTGQEMREREGILPIALAH 240
QY 241 DAIGWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVGMAI 300
DB 241 DAIGWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVGMAI 300
QY 301 DRADRDRLRLNMLFRGCPDRRYRLQRFYRLPOLIERFYAGRLTLADRLRIYVGR 360
DB 301 DRADRDRLRLNMLFRGCPDRRYRLQRFYRLPOLIERFYAGRLTLADRLRIYVGR 360
QY 361 PPIPLSQAVRCLPERPLQERA 382
DB 361 PPIPLSQAVRCLPERPLQERA 382

RESULT 2

US-09-920-923-5
Sequence 5, Application US/09920923
Publication No. US20030022273A1
GENERAL INFORMATION:
APPLICANT: Pasamontes, Luis
APPLICANT: Tsygankov, Yuri
TITLE OF INVENTION: Improved Fermentative Carotenoid Production
FILE REFERENCE: Improved Fermentative Carotenoid
CURRENT APPLICATION NUMBER: US/09/920,923
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 08/980,832
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 382
TYPE: PRT
ORGANISM: Flavobacterium sp. R1534
US-09-920-923-5

Query Match 100.0%; Score 1999; DB 10; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.1e-193;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLLIAGAGISGALLIAVDRRPPDARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARL 60
DB 1 MSHDLLIAGAGISGALLIAVDRRPPDARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARL 60

QY 61 SPIRGEMTDEVAFPDHSRRLTTGYSIEAGALIGLQGVDLRNMTHVATLDDTGATLT 120
DB 61 SPIRGEMTDEVAFPDHSRRLTTGYSIEAGALIGLQGVDLRNMTHVATLDDTGATLT 120
QY 121 DGSRIEACVIDARGAVETPHLTVGFQKFGVEIETDAPHGERPMINDATVPQMDGYRF 180
DB 121 DGSRIEACVIDARGAVETPHLTVGFQKFGVEIETDAPHGERPMINDATVPQMDGYRF 180
QY 181 IYLLPFSPTRLIEBTRYSDDGDLDDGALAQASLDYARRGWTGQEMREREGILPIALAH 240
DB 181 IYLLPFSPTRLIEBTRYSDDGDLDDGALAQASLDYARRGWTGQEMREREGILPIALAH 240
QY 241 DAIGWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVGMAI 300
DB 241 DAIGWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVGMAI 300
QY 301 DRADRDRLRLNMLFRGCPDRRYRLQRFYRLPOLIERFYAGRLTLADRLRIYVGR 360
DB 301 DRADRDRLRLNMLFRGCPDRRYRLQRFYRLPOLIERFYAGRLTLADRLRIYVGR 360
QY 361 PPIPLSQAVRCLPERPLQERA 382
DB 361 PPIPLSQAVRCLPERPLQERA 382

RESULT 3

US-10-695-980-5
Sequence 5, Application US/10695980
Publication No. US20040058410A1
GENERAL INFORMATION:
APPLICANT: Pasamontes, Luis
APPLICANT: Tsygankov, Yuri
TITLE OF INVENTION: Fermentative Carotenoid Production
FILE REFERENCE: 15464 US (C38435/125944)
CURRENT APPLICATION NUMBER: US/10/695,980
PRIOR FILING DATE: 2003-10-29
PRIOR APPLICATION NUMBER: US/09/920,923B
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 08/980,832
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 382
TYPE: PRT
ORGANISM: Flavobacterium sp. R1534
US-10-695-980-5

Query Match 100.0%; Score 1999; DB 15; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.1e-193;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLLIAGAGISGALLIAVDRRPPDARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARL 60
DB 1 MSHDLLIAGAGISGALLIAVDRRPPDARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARL 60
QY 61 SPIRGEMTDEVAFPDHSRRLTTGYSIEAGALIGLQGVDLRNMTHVATLDDTGATLT 120
DB 61 SPIRGEMTDEVAFPDHSRRLTTGYSIEAGALIGLQGVDLRNMTHVATLDDTGATLT 120
QY 121 DGSRIEACVIDARGAVETPHLTVGFQKFGVEIETDAPHGERPMINDATVPQMDGYRF 180
DB 121 DGSRIEACVIDARGAVETPHLTVGFQKFGVEIETDAPHGERPMINDATVPQMDGYRF 180
QY 181 IYLLPFSPTRLIEBTRYSDDGDLDDGALAQASLDYARRGWTGQEMREREGILPIALAH 240
DB 181 IYLLPFSPTRLIEBTRYSDDGDLDDGALAQASLDYARRGWTGQEMREREGILPIALAH 240
QY 241 DAIGWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVGMAI 300
DB 241 DAIGWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVGMAI 300

Qy 301 DEADDRFLRLNRLFRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLADRLRIVTGR 360
 Db 301 DEADDRFLRLNRLFRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLADRLRIVTGR 360
 Qy 361 PPIPLSOAVRCLPERPILQERA 382
 Db 361 PPIPLSOAVRCLPERPILQERA 382

RESULT 4

US-09-941-947A-30
 ; Sequence 30, Application US/09941947A
 ; Publication No. US20030003528A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Brzostowicz, Patricia C.
 ; APPLICANT: Cheng, Qiong
 ; APPLICANT: Dicosimo, Deana J.
 ; APPLICANT: Koffas, Mattheos
 ; APPLICANT: Miller, Edward S. Jr.
 ; APPLICANT: Odom, J. Martin
 ; APPLICANT: Picataggio, Steve
 ; APPLICANT: Rouviere, Pierre B.
 ; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
 ; FILE REFERENCE: CI.903 US NA
 ; CURRENT APPLICATION NUMBER: US/09/941,947A
 ; PRIOR FILING DATE: 2001-09-01
 ; PRIOR APPLICATION NUMBER: 60/229,907
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 60/229,858
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 30
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Pantoea stewartii
 ; US-09-941-947A-30

Query Match 39.2%; Score 783.5; DB 10; Length 382;
 Best Local Similarity 43.7%; Pred. No. 2.7e-70;
 Matches 165; Conservative 64; Mismatches 136; Indels 13; Gaps 6;

Qy 3 HLLLAGAGLGGALALAVRRDPARIVMLDARSGPSDQHTWSCDHTLS--PEWTLAR 59
 Db 5 YDLILVAGLGLANGLLALRLQOHPDWRILILIEGPGAGNHTWSTHEEDLTINQRMIA- 63
 Qy 60 LSPIRGEWTDQEVAFPDHSRRLTTGYSIEAGALIGLQ--GVDLRNTHVATLDDTG 116
 Db 64 --PLVVHMPDYQVAFQRRRHVNSGYCVTSRHFAGILRQOFGHLMHTVSAVHAES 121
 Qy 117 ATLTDGSRLEAACVTDARGAVETPHLTGFGQFVGEIETDAPHGVERPMINDATVPQND 176
 Db 122 VQLADGRILIHASTVIDRGYTPDSALRVGFOAFIQEWQLSAPHGISPIINDATVDOON 181
 Qy 177 GYRFYILPFSPTRLIEPTRYSDGDLDDGALAQSILDYARRGWTOE-WRRRGILP 235
 Db 182 GYRFYITLPSATLIEPTHTYIDKANLQAEARQINRDYARQGPLOTTLREBQALP 241
 Qy 236 IALADDAIGFWRDHAQAVPVGAGLFPVTVGSLPYAAQVADAIARDL--TTASARRA 294
 Db 242 ILTGDNQRFWQOQOQAC--SGLRAGLFPHTTGYSLPLAVALADRLSALDVTSSVHOT 299
 Qy 295 VAGMAIDRADRDRLFLNRLFRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLADRL 354
 Db 300 IAHFAQQRWQOQGFPMNLNRLFLAGPAGESRWRVMQRFYGLBEDLIARFYAGRLTVDRL 359
 Qy 355 RIVTGRPPIPLSOAVRCL 372
 Db 360 RILSGKPPVPFPAALQAI 377

RESULT 5
 US-10-218-118-6

; Sequence 6, Application US/10218118
 ; Publication No. US20030148319A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Brzostowicz, Patricia
 ; APPLICANT: Rouviere, Pierre
 ; APPLICANT: Picataggio, Stephen
 ; APPLICANT: Cheng, Qiong
 ; TITLE OF INVENTION: Genes Encoding Carotenoid Compounds
 ; FILE REFERENCE: CI.1876 US NA
 ; CURRENT APPLICATION NUMBER: US/10/218,118
 ; PRIOR FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: 60/312,646
 ; PRIOR FILING DATE: 2001-08-15
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 6
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Pantoea stewartii
 ; US-10-218-118-6

Query Match 39.2%; Score 783.5; DB 14; Length 382;
 Best Local Similarity 43.7%; Pred. No. 2.7e-70;
 Matches 165; Conservative 64; Mismatches 136; Indels 13; Gaps 6;

Qy 3 HLLLAGAGLGGALALAVRRDPARIVMLDARSGPSDQHTWSCDHTLS--PEWTLAR 59
 Db 5 YDLILVAGLGLANGLLALRLQOHPDWRILILIEGPGAGNHTWSTHEEDLTINQRMIA- 63
 Qy 60 LSPIRGEWTDQEVAFPDHSRRLTTGYSIEAGALIGLQ--GVDLRNTHVATLDDTG 116
 Db 64 --PLVVHMPDYQVAFQRRRHVNSGYCVTSRHFAGILRQOFGHLMHTVSAVHAES 121
 Qy 117 ATLTDGSRLEAACVTDARGAVETPHLTGFGQFVGEIETDAPHGVERPMINDATVPQND 176
 Db 122 VQLADGRILIHASTVIDRGYTPDSALRVGFOAFIQEWQLSAPHGISPIINDATVDOON 181
 Qy 177 GYRFYILPFSPTRLIEPTRYSDGDLDDGALAQSILDYARRGWTOE-WRRRGILP 235
 Db 182 GYRFYITLPSATLIEPTHTYIDKANLQAEARQINRDYARQGPLOTTLREBQALP 241
 Qy 236 IALADDAIGFWRDHAQAVPVGAGLFPVTVGSLPYAAQVADAIARDL--TTASARRA 294
 Db 242 ILTGDNQRFWQOQOQAC--SGLRAGLFPHTTGYSLPLAVALADRLSALDVTSSVHOT 299
 Qy 295 VAGMAIDRADRDRLFLNRLFRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLADRL 354
 Db 300 IAHFAQQRWQOQGFPMNLNRLFLAGPAGESRWRVMQRFYGLBEDLIARFYAGRLTVDRL 359
 Qy 355 RIVTGRPPIPLSOAVRCL 372
 Db 360 RILSGKPPVPFPAALQAI 377

RESULT 6
 US-10-363-567-30
 ; Sequence 30, Application US/10363567
 ; Publication No. US20040077068A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E. I. du Pont de Nemours & Company
 ; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
 ; FILE REFERENCE: CI.903 PCT
 ; CURRENT APPLICATION NUMBER: US/10/363,567
 ; PRIOR FILING DATE: 2003-09-04
 ; PRIOR APPLICATION NUMBER: US 60/229858
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: US 60/229907
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 30
 ; LENGTH: 382
 ; TYPE: PRT

```
; ORGANISM: Pantoea stewartii
US-10-567-30

Query Match      39.2%; Score 783.5; DB 15; Length 382;
Best Local Similarity 43.7%; Pred. No. 2.7e-70;
Matches 165; Conservative 64; Mismatches 136; Indels 13; Gaps 6;

QY      3 HDLLIAGAGLGGALIALAVDRRDPARIYMLDARSGPSDQHTWSCHDTDL$---PEWTLAR 59
DB      5 YDLIIVGAGLANGLIALRLQOOHPDMRILLIEAGPEAGNHTWSFHEEDLTINQHRWLA- 63
QY      60 LSPIRGEMTDOEVAFPDHSRRLTGYGSIKAGALIGLLQ---GYDLKRNTHVATLDDTG 116
DB      64 --PLVVHHPDYQVRFPPQRRRHVNSGYCVTSRHPAGILRQOFQGHLMHTAVSAVHES 121
QY      117 ATLTDGSRIEACVLDAGAVETPHLTVGFOKVGVEIETDAPHGVERPMIDATVPQMD 176
DB      122 VOLADGRITIHASTVIDGRYTPDSALRVGFQAFIGQEWQLSAPHGLSSPILMDATVPQON 181
QY      177 GYRFYTLPLSPFTRLIEDTRYSDGGDLDDGALAQASLDYARSGWTQOE-MRRRGILP 235
DB      182 GYRFYTLPLSATALLIEDTHYIDKANLQABRARQINIDYARQOWPLQTLIREQGLP 241
QY      236 IALAHDAIGFMRDHAQAVPVGLAGLPHPTGYSLPYAAQVADAIARDL-TTASARRA 294
DB      242 ITLTGDNQFQWQOQPOAC--SGLRAGLPHPTGYSLPYALVALDRLSLDVFTSSVHQT 299
QY      295 VEGMAIDRADDRLFRILNRMFLRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLARL 354
DB      300 IAHFAQQRWQOQGFPRMLNRMFLLAGPABSRWRVWQRYGPEDLIAIFYAGKLTVTDL 359
QY      355 RIVTGRPPILPSQAVRCL 372
DB      360 RILSGKPPVPVPAALQAI 377

RESULT 7
US-10-466-656-4
; Sequence 4, Application US/10466656
; Publication No. US20040078846A1
; GENERAL INFORMATION:
; APPLICANT: desouza, Mervyn L.
; APPLICANT: Schroeder, William A.
; APPLICANT: Kollmann, Sherry R.
; APPLICANT: May, Colleen A.
; TITLE OF INVENTION: Carotenoid Biosynthesis
; FILE REFERENCE: 12794-004US1
; CURRENT APPLICATION NUMBER: US/10/466,656
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/288,984
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/264,329
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Pantoea stewartii
US-10-466-656-4

Query Match      39.2%; Score 783.5; DB 15; Length 382;
Best Local Similarity 43.7%; Pred. No. 2.7e-70;
Matches 165; Conservative 64; Mismatches 136; Indels 13; Gaps 6;
```

```
QY      117 ATLTDGSRIEACVLDAGAVETPHLTVGFOKVGVEIETDAPHGVERPMIDATVPQMD 176
DB      122 VOLADGRITIHASTVIDGRYTPDSALRVGFQAFIGQEWQLSAPHGLSSPILMDATVPQON 181
QY      177 GYRFYTLPLSPFTRLIEDTRYSDGGDLDDGALAQASLDYARSGWTQOE-MRRRGILP 235
DB      182 GYRFYTLPLSATALLIEDTHYIDKANLQABRARQINIDYARQOWPLQTLIREQGLP 241
QY      236 IALAHDAIGFMRDHAQAVPVGLAGLPHPTGYSLPYAAQVADAIARDL-TTASARRA 294
DB      242 ITLTGDNQFQWQOQPOAC--SGLRAGLPHPTGYSLPYALVALDRLSLDVFTSSVHQT 299
QY      295 VEGMAIDRADDRLFRILNRMFLRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLARL 354
DB      300 IAHFAQQRWQOQGFPRMLNRMFLLAGPABSRWRVWQRYGPEDLIAIFYAGKLTVTDL 359
QY      355 RIVTGRPPILPSQAVRCL 372
DB      360 RILSGKPPVPVPAALQAI 377

RESULT 8
US-10-735-019-6
; Sequence 6, Application US/10735019
; Publication No. US20040191863A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Suh, Wonchul
; TITLE OF INVENTION: Mutations affecting Plasmid Copy Number
; FILE REFERENCE: CL2029 US NA
; CURRENT APPLICATION NUMBER: US/10/735,019
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/434973
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Pantoea stewartii
US-10-735-019-6

Query Match      39.2%; Score 783.5; DB 16; Length 382;
Best Local Similarity 43.7%; Pred. No. 2.7e-70;
Matches 165; Conservative 64; Mismatches 136; Indels 13; Gaps 6;

QY      3 HDLLIAGAGLGGALIALAVDRRDPARIYMLDARSGPSDQHTWSCHDTDL$---PEWTLAR 59
DB      5 YDLIIVGAGLANGLIALRLQOOHPDMRILLIEAGPEAGNHTWSFHEEDLTINQHRWLA- 63
QY      60 LSPIRGEMTDOEVAFPDHSRRLTGYGSIKAGALIGLLQ---GYDLKRNTHVATLDDTG 116
DB      64 --PLVVHHPDYQVRFPPQRRRHVNSGYCVTSRHPAGILRQOFQGHLMHTAVSAVHES 121
QY      117 ATLTDGSRIEACVLDAGAVETPHLTVGFOKVGVEIETDAPHGVERPMIDATVPQMD 176
DB      122 VOLADGRITIHASTVIDGRYTPDSALRVGFQAFIGQEWQLSAPHGLSSPILMDATVPQON 181
QY      177 GYRFYTLPLSPFTRLIEDTRYSDGGDLDDGALAQASLDYARSGWTQOE-MRRRGILP 235
DB      182 GYRFYTLPLSATALLIEDTHYIDKANLQABRARQINIDYARQOWPLQTLIREQGLP 241
QY      236 IALAHDAIGFMRDHAQAVPVGLAGLPHPTGYSLPYAAQVADAIARDL-TTASARRA 294
DB      242 ITLTGDNQFQWQOQPOAC--SGLRAGLPHPTGYSLPYALVALDRLSLDVFTSSVHQT 299
QY      295 VEGMAIDRADDRLFRILNRMFLRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLARL 354
DB      300 IAHFAQQRWQOQGFPRMLNRMFLLAGPABSRWRVWQRYGPEDLIAIFYAGKLTVTDL 359
QY      355 RIVTGRPPILPSQAVRCL 372
```

```

Db      360 RILSGKPPVVFALQAI 377

RESULT 9
US-10-734-778-6
; Sequence 6, Application US/10734778
; Publication No. US2004020936SAI
GENERAL INFORMATION:
APPLICANT: E.I. duPont de Nemours and Company, Inc.
APPLICANT: Suh, Wonchul
TITLE OF INVENTION: PARALLEL CHROMOSOMAL STACKING OF TRAITS IN BACTERIA
FILE REFERENCE: CL2026 US NA
CURRENT APPLICATION NUMBER: US/10/734,778
PRIOR FILING DATE: 2003-12-12
PRIOR APPLICATION NUMBER: US 60/434773
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.2
SEQ ID NO 6
LENGTH: 382
TYPE: PRT
ORGANISM: Pantoea stewartii
US-10-734-778-6

Query Match          39.2%; Score 783.5; DB 16; Length 382;
Best Local Similarity 43.7%; Pred. No. 2,7e-70;
Matches 165; Conservative 64; Mismatches 136; Indels 13; Gaps 6;

Cy      3 HLLLAGAGLGGALLAVDRBRPDARIWMLDARSGPSDOHTWSCHDTDS--PEWLAR 59
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      5 YDLILVAGAGLANGLILRLQQQHPRMRLILEACGEPAGCNHTWSFHEEDLTINQRMTA- 63
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Cy      60 LSPIRGEWTDOEVAEPFDHSRRLLTTGYGSIEAGALIGLQ---GVDLRMNTHVALTDPTG 116
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db      64 --PLVVHHPDVQVRFPQRRRHVNSGYCVTSRHFAGILLQFGQHLMLHTAVASVNAES 121
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Cy      117 ATLTDGSRLEAACVIDARGAVETPHLTVGFQKFVGVEIETDAPHGVEREMTADATVPMD 176
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db      122 VQLADGRILHASTVIDRGRTTPDSALRVGFQAFIGEQWQSAPHGSLSSPIIMDAIVDQN 181
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Cy      177 GYRFYLLEPSPFRILLIEDTRYSDGDGLDDGALAASLDVAARCGWNGOE-MREERGLIP 235
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db      182 GRPFYVTLPLSATALLIEDHYIDKANLQERARQNRIDVAARCGWPLQTLLREEQGLP 241
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Cy      236 IMLADAIFPMWRDHAGAVPVGLAGLFHPVTGYSLPYAQAVADAIARDL-TTASARA 294
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db      242 ITLTGDNRPFWCQQOQPAC--SGLRAGLFHFTTGSLPLAVALADRLSALDVFTSSVHQ 299
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Cy      295 VEGMAIDRADRDRLFLNNLMFLRCGPDDRYYRLLOFPYRIPOPLIERFYAGRLLTLADR 354
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db      300 IAHFAQQRWQQQGFPMNLNMLFLAGABSRMRVMQGFGYLEDLIARFAYAKLTVTDRL 359
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Cy      355 RIIVGRRPIPLSQAVRCL 372
        :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db      360 RILSGKPPVVFALQAI 377

RESULT 10
US-10-735-442-6
; Sequence 6, Application US/10735442
; Publication No. US20040219629AI
GENERAL INFORMATION:
APPLICANT: E. I. duPont de Nemours and Company, Inc.
APPLICANT: Suh, Wonchul
APPLICANT: Rouviere, Pierre
TITLE OF INVENTION: Increasing Carotenoid Production in Bacteria Via Chromosomal
FILE REFERENCE: CL2027 US NA
CURRENT APPLICATION NUMBER: US/10/735,442
PRIOR FILING DATE: 2003-12-12
PRIOR APPLICATION NUMBER: US 60/434618

```

[illegible]

```

Oy 3DILLIAGAGLSGILALAAVDRRRDARIWLDARSQSDHWTSCHTDLS---PEWIR 59
Db 5YDLILVAGELANGLILALRLOQOHPRDKRILILEGPEAGSNHTWSFHEEDLILNOHWTA- 63
Oy 60ISPIRGEWTDQVEAFPPDSRRRLTTGYSIEAGALLILO---GVDIRMWNHVALTLDDTG 116
Db 64--PLVYHHPRDVQYRFPQRRRHVNSGYUUCVTSNHFACILINQOQHMLHTAVSAVNAES 121
Oy 117ATLTGDSRIEACVIDARGAVETPHLTVGFOKVGVIEIFTDAPRGVERPMIMDAITVPOMD 176
Db 122VOLAGRIIHASTVIDIGRGYTPDSALRVGFOAFIGQWOLSAFGLSSPIIMDTVTQON 181
Oy 177GYRFYILPEFPRIILIEDTRYSDGGDLDDGALQAASLDVAARGMWGOE-MRAERGITLP 235
Db 182GRFYFTLPSATALLIEDTHYTDKALQAEARARONIRDTAARGWPLQTLIRREQALP 241
Oy 236IALAHDAIEFWMDHAGAVPVGLGAGLFHFPVTGSLPYAAQVADAIARBL-TRASARA 294
Db 242ITLTGNRQFPWQOQPOAC--SGLRAGLFHHTTGSPLALVALADRLSALDVFSTSSVHOT 299
Oy 295VAGMAIDRADRDREFTLNNRMLFRGCPBDBRRYRILLOFYRLPOPLIERFYAGRLTLADRL 354
Db 300IAHFAQORWQOQGFPMILNRMFLAGBAESRMWRMVGFGHPEBLLIARFYAGKLTVDRL 359
Oy 355RIVTGRPPILPSQAVRCL 372
Db 360RIISGRPVVFVAALQAI 377

```

```

RESULT 12
US-10-866-906-6
? Sequence 6, Application US/10866906
? Publication No. US20050019852A1
? GENERAL INFORMATION:
? APPLICANT: Suh, Wonchul
? APPLICANT: Rouviere, Pierre
? APPLICANT: Cheng, Qiong
? APPLICANT: Tao, Luan
? TITLE OF INVENTION: Production of Aromatic Carotenoids in Gram Negative Bacteria
? FILE REFERENCE: CL2128 US NA
? CURRENT APPLICATION NUMBER: US/10/866,906
? PRIOR FILING DATE: 2004-07-08
? PRIOR APPLICATION NUMBER: US 60/486,106
? PRIOR FILING DATE: 2003-07-10
? NUMBER OF SEQ ID NOS: 57
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 6
? LENGTH: 382
? TYPE: PRT
? ORGANISM: Pantoea stewartii
US-10-866-906-6

Query Match          39.2%; Score 783.5; DB 17; Length 382;
Best Local Similarity 43.7%; Pred. No. 2,7e-70;
Matches   165; Conservative    64; Mismatches 136; Indels   13; Gaps     6;

OY      3 HDLLIAGAGLSCALLALAVDRRRDARIIVMLDARSGPSDDHTWSCHDTDS---PEWILAR 59
       :|::||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      5 YDLIIWGAGIAMLIILRLQOOHPDMRILLIEAGPEAGGNHNTVSFHEEDLTINOHRWIA- 63
       ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::

OY      60 ISPIRGEMTDDEOVAPPPDHRSRLTTTGYSLEAGLIGILO--GVDIRMWNHVATLDITDG 116
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      64 --PLVVHHMPDYVVRFPQRKRVNNSGYCVTSNRFAGLIKFOFHLMTLTAVASVAHAES 121
       |::||::||::||::||::||::||::||::||::||::||::||::||::||::||::

OY      117 ATLTDGSRIEACVIDARGAVETPHLVGFQKFVGVEIETDPAGVERPMIMDAITVPQMID 176
        |::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      122 VOLADGRILIHASTVIIDRGYTPPSALKRVGFQAFTGWQSASPRGLSSPIIMDATVDQN 181
         |::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

OY      177 GVRFYLIPSPKRILLIEDTRYSDGDGLDDGALAASAQLDVAAARPGMWTGE-MREERGILT 235
        ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      182 GRPFVTLPISATALIIDETHYDKNALQERRARNONTRIDAARGMWPLQTLLIREQGKL 241
         ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

OY      236 IALAHAI GFWRDHGAOGAVNGVLGAGLFHFPTGTSLFEPAOVAADAIALARDL-TTASARA 294
```

Db 242 ITLTGNGRQWQOOPAC--SSLRAGLTHPTTGSPLPAVALADRSLADLVFSSVHOT 239

Qy 295 VRCMAIDRADRDREFLINRMFLRCPCPDRRRLIQRFLYLPQULIERPFAIGRTTLADR 354

Db 300 IAHFAOQRMQOQOQFFFMILNRMFLTAGPESRWRVQRGYGLPEDLIARFAGKLTVDRL 359

Qy 355 RIVGRPPIPLSQAVRC 372

Db 360 RIISGRPEVVFPAALQAT 377

```

RESULT 13
US-10-808-807-6
; Sequence 6, Application US/10808807
; Publication No. US2004025363A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Co., Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Juan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2365 US NA
; CURRENT APPLICATION NUMBER: US/10/808,807
; CURRENT FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/477,874
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pantoea agglomerans strain DC404
US-10-808-807-6

Query Match      38.3%; Score 766; DB 16; Length 389;
Best Local Similarity 44.7%; Pred. No. 1,6e-68;
Matches 169; Conservative 62; Mismatches 133; Indels 14; Gaps 9

QY 4 DLTLAGAGLGGATLALAVDRRPPARIYIMLARSGPSQDHTWSCDPTLSPFWLARLSP 63
DB 5 DLTIVGAGLGGATLAWRKQHPPTLAVLMLECGDAPGSGNHTWSFHQDITTPAQHMLAPL 64
QY 64 RRGWTDQEVAFPPDHSRRLTTGYGSI---AGALIGLQGVDRMNTHTVATLDDTGATL 119
DB 65 VAHRDGDYVHPNNVSRLLTHDGYLLTSTRFAQARGLMK-ENLLTNTVTSVSGQEVTL 123
QY 120 TDGSRILEACYIDARGAETPHLYTGFGKFGVEIETDAPHCVEEPMLMDATVPQMDGR 179
DB 124 SDGRFTTGAAYIDGXYSPHLSIGXYQAFIGQEWQLTAPHGLTPIILMDARVAAGNGYR 183
QY 180 FYLLPSPPTLLIEDTYSDDGDDID-DGALAAQASLDYAAHRCW-TGQEMRRERGIPLIA 237
DB 184 FVYTLPLSADTLLIEDTYIDGPTLDAASAPRRIA-DARQGGWOLARLVREBQALPIT 242
QY 238 LAHDAIGFWRD-HAOGAPV-GLAGLPHYVYGSYPYAAQVADAIA-ARDITTSARBA 294
DB 243 LSGDPAALFHQFHHQ--PVSGLRAGLPHATTGYSLPLAVNLADRIANAPGLHQAALYQL 299
QY 295 VRGWAIDRADRFLRLINRLFRGCCPPDRRYRLQRFYRLPQPLIERPYAGRLTLADR 354
DB 300 IADFAARHMQTORFFRILINRLFLAGTDDQWRVQRVFGYQDEQLIARFYAQLRSADBA 359
QY 355 RIVTGRPIPLSQAVRCL 372
DB 360 RLILGKPPVPIVGAIKAL 377

RESULT 14
US-10-810-733-8
; Sequence 8, Application US/10810733

```

Query Match	38.0%;	Score 760;	DB 17;	Length 387;
Best Local Similarity	44.0%;	Pred. No. 6.6e-68;		
Matches 168;	Conservative 54;	Mismatches 146;	Indels 14;	Gaps 6;

RESULT 15
US-10-804-677-6

Sequence 6, Application US/10804677
Publication No. US2004022483A1
GENERAL INFORMATION:
APPLICANT: E.I. duPont de Nemours and Company, Inc.
APPLICANT: Chemt; Qiong
APPLICANT: Tao; Luan
APPLICANT: Sedkova, Natalia
TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
FILE REFERENCE: CL2346 US NA
CURRENT APPLICATION NUMBER: US/10/804,677
CURRENT FILING DATE: 2004-03-19
PRIOR APPLICATION NUMBER: US 60/468,596
PRIOR FILING DATE: 2003-05-07
PRIOR APPLICATION NUMBER: US 60/527,083
PRIOR FILING DATE: 2003-12-03
NUMBER OF SEQ ID NOS: 20

Query Match	38.0%;	Score 759;	DB 16;	Length 388;
Best Local Similarity	43.2%;	Pred. NO. 8.4e-68;		
Matches 163;	Conservative 61;	Mismatches 143;	Indels 10;	Gaps 5;

Search completed: June 17, 2005, 19:55:52
Job time : 132.394 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 19:21:35 ; Search time 37.3429 Seconds
(without alignments)
763.623 Million cell updates/sec

Title: US-10-695-980-5

Perfect score: 1999

Sequence: 1 MSHDLLAGAGLSGALIALA.....IPLSQAVRCLEPRLQERA 382

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/ECTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	382	3	US-08-660-645A-7
2	1999	100.0	382	3	US-09-298-718-7
3	1999	100.0	382	3	US-09-546-969-7
4	1999	100.0	382	3	US-08-980-832-5
5	1999	100.0	382	4	US-09-547-267-7
6	1999	100.0	382	4	US-09-920-923B-5
7	1452	72.6	386	2	US-08-663-310-6
8	1452	72.6	386	2	US-09-006-491-6
9	1452	72.6	386	1	US-09-335-919-6
10	810.5	40.5	382	1	US-07-783-705A-3
11	713.5	35.7	374	1	US-08-095-726-14
12	713.5	35.7	374	1	US-08-096-623A-14
13	195.5	9.8	511	4	US-09-323-998B-55
14	170.5	8.5	498	4	US-09-323-998B-57
15	163.5	8.2	502	4	US-09-323-998B-56
16	161.5	8.1	500	4	US-09-323-998B-58
17	161	8.1	500	4	US-09-323-998B-59
18	158	7.9	498	3	US-09-134-607A-19
19	157.5	7.9	511	4	US-09-201-641-2
20	157.5	7.9	511	4	US-09-323-998B-60
21	153	7.7	498	3	US-09-134-607A-17
22	153	7.7	498	3	US-09-134-607A-18
23	137	6.9	503	4	US-09-323-998B-61
24	132	6.6	498	2	US-08-702-598-2
25	127	6.4	456	4	US-08-624-125-20
26	127	6.4	456	4	US-08-937-155-20
27	126.5	6.3	526	4	US-09-323-998B-53

28	123.5	6.2	411	1	US-08-399-561-2	Sequence 2, Appl
29	122	6.1	524	1	US-08-624-125-2	Sequence 2, Appl
30	122	6.1	524	1	US-08-624-125-21	Sequence 21, Appl
31	122	6.1	524	4	US-08-937-155-2	Sequence 2, Appl
32	122	6.1	524	4	US-08-937-155-21	Sequence 21, Appl
33	122	6.1	524	4	US-09-323-998B-2	Sequence 2, Appl
34	122	6.1	524	4	US-09-323-998B-49	Sequence 49, Appl
35	119	6.0	524	4	US-09-323-998B-21	Sequence 21, Appl
36	118.5	5.9	533	4	US-09-323-998B-26	Sequence 26, Appl
37	117.5	5.9	487	4	US-09-252-991A-21980	Sequence 21980, A
38	113	5.7	529	4	US-09-323-998B-23	Sequence 23, Appl
39	113	5.7	529	4	US-09-323-998B-47	Sequence 47, Appl
40	113	5.7	529	4	US-09-323-998B-50	Sequence 50, Appl
41	112	5.6	378	4	US-09-323-998B-25	Sequence 25, Appl
42	112	5.6	378	4	US-09-323-998B-48	Sequence 48, Appl
43	112	5.6	426	4	US-09-252-991A-27522	Sequence 27522, A
44	111	5.6	374	4	US-09-323-998B-27	Sequence 27, Appl
45	111	5.6	516	3	US-09-201-641-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-660-645A-7

; Sequence 7, Application US/08660645A

; Patent No. 6087152

; GENERAL INFORMATION:

; APPLICANT: Hoffmann, Hans-Peter

; APPLICANT: Pasamontes, Luis

; APPLICANT: Tessier, Michel

; APPLICANT: van Loon, Adolphus

; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION

; NUMBER OF SEQUENCES: 47

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann-La Roche Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: NJ

; COUNTRY: USA

; ZIP: 07110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/660,645A

; FILING DATE: 07-JUN-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 95108888.9

; FILING DATE: 09-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Pokras, Bruce A.

; REGISTRATION NUMBER: 32,748

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-5801

; TELEFAX: (201) 235-2363

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 382 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-660-645A-7

Query Match 100.0%; Score 1999; DB 3; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.56-211;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLLIAGAGISGALLAVALVDRPPDARIYVMDARSGPSDQHTWSCHDTLSPFWLRL 60
DB 1 MSHDLLIAGAGISGALLAVALVDRPPDARIYVMDARSGPSDQHTWSCHDTLSPFWLRL 60
QY 61 SPIRGEMTDOEVAEPDHSRRLTTGYSTEGALIGLLQGVDLRNTHVATLDDTGATLT 120
DB 61 SPIRGEMTDOEVAEPDHSRRLTTGYSTEGALIGLLQGVDLRNTHVATLDDTGATLT 120
QY 121 DGSRIEACVIDARGAVETPHLTGQFQVGEIETDAPHGVERPMINDATVPQMDGYRF 180
DB 121 DGSRIEACVIDARGAVETPHLTGQFQVGEIETDAPHGVERPMINDATVPQMDGYRF 180
QY 181 IYLLPFSPTRLIIEBTRYSDDGDLDDGALAQASLDYAARRGWTQEMREREGILPIALAH 240
DB 181 IYLLPFSPTRLIIEBTRYSDDGDLDDGALAQASLDYAARRGWTQEMREREGILPIALAH 240
QY 241 DAIGFWRHQAQAVPVGAGLFPHTVGSLEPYAAQVADALAAADLTASARRAVGMAI 300
DB 241 DAIGFWRHQAQAVPVGAGLFPHTVGSLEPYAAQVADALAAADLTASARRAVGMAI 300
QY 301 DRADRDRLRLNRLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIYVGR 360
DB 301 DRADRDRLRLNRLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIYVGR 360
QY 361 PPIPLSQAVRCLPERPILQERA 382
DB 361 PPIPLSQAVRCLPERPILQERA 382

RESULT 2
US-09-298-718-7
; Sequence 7, Application US/09298718
; Patent No. 6124113

GENERAL INFORMATION:

APPLICANT: Hohmann, Hans-Peter
APPLICANT: Pasamontes, Luis
APPLICANT: Tessier, Michel
APPLICANT: van Loon, Adolpheus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/298,718
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: RAN 6002/170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-298-718-7
Query Match 100.0%; Score 1999; DB 3; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.5e-211;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLLIAGAGISGALLAVALVDRPPDARIYVMDARSGPSDQHTWSCHDTLSPFWLRL 60
DB 1 MSHDLLIAGAGISGALLAVALVDRPPDARIYVMDARSGPSDQHTWSCHDTLSPFWLRL 60
QY 61 SPIRGEMTDOEVAEPDHSRRLTTGYSTEGALIGLLQGVDLRNTHVATLDDTGATLT 120
DB 61 SPIRGEMTDOEVAEPDHSRRLTTGYSTEGALIGLLQGVDLRNTHVATLDDTGATLT 120
QY 121 DGSRIEACVIDARGAVETPHLTGQFQVGEIETDAPHGVERPMINDATVPQMDGYRF 180
DB 121 DGSRIEACVIDARGAVETPHLTGQFQVGEIETDAPHGVERPMINDATVPQMDGYRF 180
QY 181 IYLLPFSPTRLIIEBTRYSDDGDLDDGALAQASLDYAARRGWTQEMREREGILPIALAH 240
DB 181 IYLLPFSPTRLIIEBTRYSDDGDLDDGALAQASLDYAARRGWTQEMREREGILPIALAH 240
QY 241 DAIGFWRHQAQAVPVGAGLFPHTVGSLEPYAAQVADALAAADLTASARRAVGMAI 300
DB 241 DAIGFWRHQAQAVPVGAGLFPHTVGSLEPYAAQVADALAAADLTASARRAVGMAI 300
QY 301 DRADRDRLRLNRLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIYVGR 360
DB 301 DRADRDRLRLNRLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIYVGR 360
QY 361 PPIPLSQAVRCLPERPILQERA 382
DB 361 PPIPLSQAVRCLPERPILQERA 382

RESULT 3
US-09-546-969-7
; Sequence 7, Application US/09546969
; Patent No. 6207409

GENERAL INFORMATION:

APPLICANT: Hohmann, Hans-Peter
APPLICANT: Pasamontes, Luis
APPLICANT: Tessier, Michel
APPLICANT: van Loon, Adolpheus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/546,969
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: RAN 6002/170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-546-969-7

Query Match 100.0%; Score 1999; DB 3; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.5e-211;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLIIAGAGSALIALAVDRRDPDARIWMDARSGSDQHTWSCHDTDISPEWLAHL 60
DB 1 MSHDLIIAGAGSALIALAVDRRDPDARIWMDARSGSDQHTWSCHDTDISPEWLAHL 60
QY 61 SPIRGEWTDQEVAFPDHRRRLTTGYSIEAGALIGLQGVDRMNTHTVATLDDTGATLT 120
DB 61 SPIRGEWTDQEVAFPDHRRRLTTGYSIEAGALIGLQGVDRMNTHTVATLDDTGATLT 120
QY 121 DGSRIEACVIDARGAVETPHLVGFQKVEIETDAPHGVERPMIMDATVPQMDGYRF 180
DB 121 DGSRIEACVIDARGAVETPHLVGFQKVEIETDAPHGVERPMIMDATVPQMDGYRF 180
QY 181 IYLLPSPTRILLIEDTRYSDGDLDDGALAQAQLDYAARGWTGQEMREREGILPIALAH 240
DB 181 IYLLPSPTRILLIEDTRYSDGDLDDGALAQAQLDYAARGWTGQEMREREGILPIALAH 240
QY 241 DAIGFWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVAGMAI 300
DB 241 DAIGFWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVAGMAI 300
QY 301 DRADDRFLRLNRLMFRGCPDRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360
DB 301 DRADDRFLRLNRLMFRGCPDRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360
QY 361 PPIPLSOAVRCLPERPILQERA 382
DB 361 PPIPLSOAVRCLPERPILQERA 382

RESULT 4
US-08-980-832-5
Sequence 5, Application US/08980832B
Patent No. 6231204
GENERAL INFORMATION:
APPLICANT: Pasamontes, Luis
APPLICANT: Teyganikov, Yuri
TITLE OF INVENTION: Improved Fermentative Carotenoid Production
FILE REFERENCE: Improved Fermentative Carotenoid Production
CURRENT APPLICATION NUMBER: US/08/980,832B
CURRENT FILING DATE: 1997-12-01
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 382
TYPE: PRT
ORGANISM: Flavobacterium sp. R1534
US-08-980-832-5

Query Match 100.0%; Score 1999; DB 3; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.5e-211;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSHDLIIAGAGSALIALAVDRRDPDARIWMDARSGSDQHTWSCHDTDISPEWLAHL 60
DB 1 MSHDLIIAGAGSALIALAVDRRDPDARIWMDARSGSDQHTWSCHDTDISPEWLAHL 60
QY 61 SPIRGEWTDQEVAFPDHRRRLTTGYSIEAGALIGLQGVDRMNTHTVATLDDTGATLT 120
DB 61 SPIRGEWTDQEVAFPDHRRRLTTGYSIEAGALIGLQGVDRMNTHTVATLDDTGATLT 120
QY 121 DGSRIEACVIDARGAVETPHLVGFQKVEIETDAPHGVERPMIMDATVPQMDGYRF 180

DB 121 DGSRIEACVIDARGAVETPHLVGFQKVEIETDAPHGVERPMIMDATVPQMDGYRF 180
QY 181 IYLLPSPTRILLIEDTRYSDGDLDDGALAQAQLDYAARGWTGQEMREREGILPIALAH 240
DB 181 IYLLPSPTRILLIEDTRYSDGDLDDGALAQAQLDYAARGWTGQEMREREGILPIALAH 240
QY 241 DAIGFWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVAGMAI 300
DB 241 DAIGFWRDHAQAVPVGAGLFHPVTGYSLPYAAQVADAIARDLTTASARRAVAGMAI 300
QY 301 DRADDRFLRLNRLMFRGCPDRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360
DB 301 DRADDRFLRLNRLMFRGCPDRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360
QY 361 PPIPLSOAVRCLPERPILQERA 382
DB 361 PPIPLSOAVRCLPERPILQERA 382

RESULT 5
US-09-547-267-7
Sequence 7, Application US/09547267
Patent No. 6613543
GENERAL INFORMATION:
APPLICANT: Hohmann, Hans-Peter
APPLICANT: Pasamontes, Luis
APPLICANT: Teesler, Michel
APPLICANT: van Loon, Adolpheus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/547,267
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: RAN 6002/170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-547-267-7

Query Match 100.0%; Score 1999; DB 4; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.5e-211;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSHDLIIAGAGSALIALAVDRRDPDARIWMDARSGSDQHTWSCHDTDISPEWLAHL 60
DB 1 MSHDLIIAGAGSALIALAVDRRDPDARIWMDARSGSDQHTWSCHDTDISPEWLAHL 60

Qy	6	SP1RRGEMTQOEVAFPDHSRRLLTTGGSIEAALLGLLQGVDLRNNTVATLDDTGATLT	120
Db	61	SP1RRGEMTQOEVAFPDHSRRLLTTGGSIEAALLGLLQGVDLRNNTVATLDDTGATLT	120
Qy	121	DGSRIEAACVADARGAVETTPHLTVGQFQKVGVEITDAPHGERPMIDATVPOMDGYR	180
Db	121	DGSRIEAACVADARGAVETTPHLTVGQFQKVGVEITDAPHGERPMIDATVPOMDGYR	180
Qy	181	IYLLPSPRILILEDTRYSDDGDLDDGALAQSLDYAARRCMTQGMKRRERGIIPALAH	240
Db	181	IYLLPSPRILILEDTRYSDDGDLDDGALAQSLDYAARRCMTQGMKRRERGIIPALAH	240
Qy	241	DA1GFWMEDAQGVAPVGLGAGLFHPPTGYSLEPYAAQVADALAAEDLTASARRAVRGMAI	300
Db	241	DA1GFWMEDAQGVAPVGLGAGLFHPPTGYSLEPYAAQVADALAAEDLTASARRAVRGMAI	300
Qy	301	DRADRRORFLNLRMLFRGCPDRPRVRLIQRYFLPOPLIERFYAGRLTLADRIRYTVGR	360
Db	301	DRADRRORFLNLRMLFRGCPDRPRVRLIQRYFLPOPLIERFYAGRLTLADRIRYTVGR	360
Qy	361	PR1PLSOAVRCLPERPULOERA 382	
Db	361	PR1PLSOAVRCLPERPULOERA 382	

RESULT 6
 US-09-920-923B-5
 Sequence 5, Application US/09920923B
 Patent No. 6677134
 GENERAL INFORMATION:
 APPLICANT: Paasamontes, Luis
 APPLICANT: Tsygan'kov, Yuri
 TITLE OF INVENTION: Fermentative Carotenoid Production
 FILE REFERENCE: 15464 US (C38435/125944)
 CURRENT APPLICATION NUMBER: US/09/920,923B
 PRIOR FILING DATE: 2001-08-02
 PRIOR APPLICATION NUMBER: 08/980,832
 NUMBER OF SEQ ID NOS: 66
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 5
 LENGTH: 362
 TYPE: PRT
 ORGANISM: Flavobacterium sp. R1534
 US-09-920-923B-5

Query Match	100.0%;	Score 1999;	DB 4;	Length 382;
Best Local Similarity	100.0%;	Pred. No. 2.5e-211;		
Matches 382;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1	MSHDLILAAGLSGALIALAVDRPRDAIIVWLDRSGSPSOHTWSCHDIDLSSEWIARL	60
Dh	1	MSHDLILAAGLSGALIALAVDRPRDAIIVWLDRSGSPSOHTWSCHDIDLSSEWIARL	60
Qy	61	SPIRGEMTDQEVAFPDHRSRLTTGSGISEAGALIGLLQGVDLRWHTVATLDTGATLT	120
Dh	61	SPIRGEMTDQEVAFPDHRSRLTTGSGISEAGALIGLLQGVDLRWHTVATLDTGATLT	120
Qy	121	DGSRIEAACVLDARGAVETPHLTVGQFQKVGVEIETDPHGVEMIDATVPQMDGYRF	180
Dh	121	DGSRIEAACVLDARGAVETPHLTVGQFQKVGVEIETDPHGVEMIDATVPQMDGYRF	180
Qy	181	IYLLPSPRIILLIEDPRYSDGGDLDDGALAQASLDYAARRGTOGEMRREGILPIALAH	240
Dh	181	IYLLPSPRIILLIEDPRYSDGGDLDDGALAQASLDYAARRGTOGEMRREGILPIALAH	240
Qy	241	DAIGFMRDHOQAVPVGLGAFPHPYTGSLPYAAQVADALAARLDTTASRRARVGMAL	300
Dh	241	DAIGFMRDHOQAVPVGLGAFPHPYTGSLPYAAQVADALAARLDTTASRRARVGMAL	300
Qy	301	DRADRFELLLNRMLFRGCPDPRAYLLQRFYRLPOLLERFYAGRLTLADRLLRIVTGR	360

Db 301 DRDRDRFRRLNRLMFRGCPDRDRYRLQLPQLIBRYAGRLTLADRLRYVTGR 360

QY 361 PPPLSQAVRCLEPERPLQERA 382

Db 361 PPPLSQAVRCLEPERPLQERA 382

RESULT 7
US-08-663-310-6

; Sequence 6, Application US/08663310
; Patent No. 5811273

GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5811273ihiko

;
;
; APPLICANT: KONDO, Keiji
; APPLICANT: KAWIWARA, Susumu

APPLICANT: YOKOYAMA, Akihiro
TITLE OF INVENTION: DNA STRANDS USED

TITLE OF INVENTION:	XANTHOPHYLLS AND
TITLE OF INVENTION:	XANTHOPHYLLS

```

;
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite

CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20007-5109

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;

```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0,  
; CURRENT APPLICATION DATA:  
; PATENT NUMBER: 7,689,000
```

APPLCATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: WO PCT/JP94/0
 ; FILING DATE: 26-DEC-1994
 ;

;; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-235917

```

; FILING DATE: 05-SEP-1994
; PRIOR APPLICATION DATA:
;

```

APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 00 560

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/10
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/10

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5300

TELEFAX: (202) 672-5399
TELEX: 904136
TELEPHONE FOR INFO: 202-672-5399

```

; INFORMATION FOR SEQ ID NO: 6:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 305

```

```

;
; LENGTH: 386 amino acids
;
; TYPE: amino acid
;
; MODORCV: 1.1.2.2

```

```

;          TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          NO OF CRYSTALS: 1

```

OS-08-663-310-6

Query Match	72.6%;	score 1
Best Local Similarity	71.5%;	Pred. N

Matches 276; Conservative 40; Mism

```

1 MSHDLLAGAGLSGALIALAVRRKRP
::|::|::|::|::|::|::|::|

```

Db 1 VTHDVLLAGAGLGLIALTRARP

```

61 SPIRGEMTDQEVAFPDHSRLTYGCG
|::| | | | | | | | | | | |

```

DB 61 KPLRRANWPDQEVRFPRHARRLATIGYG

301 DRADRDRLFLNLNRMFLFGCGPPDRRRLRLQRRYRLPGLIERPFVAGRLTLADRLRIYTVGR 366
361 PPIPLSQAVRCLPERPILQERA 382
|||||
361 PPIPLSQAVRCLPERPILQERA 382
|||||

LT 7
Sequence 6, Application US/08663310-6
ent No. 5811273
GENERAL INFORMATION:
APPLICANT: MISAWA, No. 5811273jhiko
APPLICANT: KONDO, Keiji
APPLICANT: KAWIWARA, Suunuu
APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
TITLE OF INVENTION: XANTHOPHYLLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington

ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
APPLICATION DATA:
PRIOR APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:

```

; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-663-310-6

Query Match          72.6%; Score 1452; DB 2; Length 386;
Best Local Similarity 71.5%; Pred. No. 4,4e-151;
Matches 276; Conservative 40; Mismatches 66; Indels 4; Gaps 2

Qy      1 MSHELLIIAGAGISGALIALAVDRRRDPARIYMLDARSGPSDFQTWSCHDTDLSPFWIARL 60
Db      1 VTHDIVLIAGAGLANGNIALALRAAPRDLRVLLDHAAGPSDGHWTSCHDPLSDMWIARL   60

Qy      61 SFRREBWTDVAFPPDSRLRTTGYGSIEAGALLGLL--CGVDLRNTWATLTDDTGAT 118
Db      61 KPIRRNNWPDEVRPRHARRLATYSGSLDGAAALADAVNRGAERIRWDSDIALDAGAT 120
```


FILING DATE: 05-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29, 768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 386 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-335-919-6

Query Match 72.6%; Score 1452; DB 3; Length 386;
Best Local Similarity 71.5%; Pred. No. 4.4e-151;
Matches 276; Conservative 40; Mismatches 66; Indels 4; Gaps 2;

QY 1 MSHDLLIAGAGISGALIALAVDRRDPARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARL 60
DB 1 VTHDVILAGAGLANGILIALRAARPDRLVLLDHAAGPSDQHTWSCDHTDLSPEWMLARL 60
QY 61 SPIRGERTDQVAFPDHSRLTTGYGSIKAGALIGLL--QGVDLKMNTHVATLDDTGAT 118
DB 61 KFLRRAMPDQEVPRPHARRLATGSGIDGALLADAVVRSGAEIRMSDIALDLAQAT 120
QY 119 LTDGRIEACVTDARAVETPHLYTGFOKFYVEIETDAPRGVRRPMIMDATTVPQMDGY 178
DB 121 LSCGRIRIAGVADRGAPSRHLYTGFOKFYVEIETDAPRGVRRPMIMDATTVPQMDGY 180
QY 179 RFTYLLPSPPTRIILIEDTRYSDGDLDDGALLAASLDYVARRGWTQENRRERGIPTAL 238
DB 181 RFTYLLPSPPTRIILIEDTRYSDGDLDDGALLAASLDYVARRGWTQENRRERGIPTAL 240
QY 239 AHDAIGFWRDHAOGAVPVGLGAGLPHPTVGYSLPYAAQVADVAIAARDLTTASARRAVR 296
DB 241 AHDAIGFWRDHAOGAVPVGLGAGLPHPTVGYSLPYAAQVADVAIAARDLTTASARRAVR 300
QY 297 GVAIRADDRFLRLNRLFRGCPDRRYRLQRFYRLPOPILIRFYAGRLTLADRLRI 356
DB 301 DYAIRADDRFLRLNRLFRGCPDRRYRLQRFYRLPOPILIRFYAGRLTLADRLRI 360
QY 357 VTGRPIPLSOAVRCLPERPILQERA 382
DB 361 VTGRPIPLSOAVRCLPERPILQENA 386

RESULT 10
US-07-783-705A-3
Sequence 3, Application US/07783705A
Patent No. 5429939
GENERAL INFORMATION:
APPLICANT: Misawa, No. 5429939ihiko
APPLICANT: Kobayashi, Kazuo
APPLICANT: Nakamura, Katsumi
APPLICANT: Yamano, Shigeyuki
TITLE OF INVENTION: DNA SEQUENCES USEFUL FOR THE
TITLE OF INVENTION: SYNTHESIS OF CAROTENOIDS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladax & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 720Kb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: N/A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07783, 705A
FILING DATE: 19911023
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1-103078
FILING DATE: 21-APR-1989
APPLICATION NUMBER: JP 2-53225
FILING DATE: 05-MAR-1990
APPLICATION NUMBER: US 07/519, 011
FILING DATE: 19-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Schwadron, Janet I.
REGISTRATION NUMBER: 33,778
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-708-1935
TELEFAX: 212-246-5959
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-783-705A-3

Query Match 40.5%; Score 810.5; DB 1; Length 382;
Best Local Similarity 45.6%; Pred. No. 1.9e-80;
Matches 171; Conservative 59; Mismatches 138; Indels 7; Gaps 4;

QY 3 HDLIIAGAGISGALIALAVDRRDPARIYMLDARSGPSDQHTWSCDHTDLSPEWMLARLSP 62
DB 5 YDLIIAGAGLANGILIALRLQOQOPMRLILLDAPRQAGNTHSHDDLLRESQHWIAP 64
QY 63 IIRGEMTQEVAFPDHSRLTTGYGSIKAGALIGLLQ--QGVDLKMNTHVATLDDTGATL 119
DB 65 LVVHHMPDYQVRFPRRRKLSNGYFCITSQRFAYVLQRFPHLMMDYAVLVNVAESVRL 124
QY 120 TDGSRIEACVTDARAVETPHLYTGFOKFYVEIETDAPRGVRRPMIMDATTVPQMDGYR 179
DB 125 KKGQVIGRAVYIDGAVANSALSVGFQAFIQQEWRLSHPLGLSSPTIMDATTVPQMDGYR 184
QY 180 RFTYLLPSPPTRIILIEDTRYSDGDLDDGALLAASLDYVARRGWTQENRRERGIPTAL 238
DB 185 RFTYLLPSPPTRIILIEDTRYSDGDLDDGALLAASLDYVARRGWTQENRRERGIPTAL 244
QY 239 AHDAIGFWRDHAOGAVPVGLGAGLPHPTVGYSLPYAAQVADVAIAARDLTTASARRAVR 297
DB 245 SGNDAPFQOQRPILAC--SGLRAGLPHPTVGYSLPYAAQVADVAIAARDLTTASARRAVR 302
QY 298 WAIDRADDRFLRLNRLFRGCPDRRYRLQRFYRLPOPILIRFYAGRLTLADRLRI 357
DB 303 FARERMOOQGFPRMLNRLFLAGPADSRWQRFGYLEDLIAFYAGRLTLADRLRI 362
QY 358 VTGRPIPLSOAVRCL 372
DB 363 SGKPPVYLAALQAI 377

RESULT 11
US-08-095-726-14
Sequence 14, Application US/08095726
Patent No. 5530188
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L
APPLICANT: Brinkhaus, Friedhelm L
APPLICANT: Mukharji, Indrani
APPLICANT: Profitic, John H
APPLICANT: Yarger, James G
APPLICANT: Yen, Huel-Che B

```

/ TITLE OF INVENTION: Beta-Carotene Biosynthesis in
/ TITLE OF INVENTION: Genetically Engineered Hosts
/ NUMBER OF SEQUENCES: 79
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Amoco Corp., Patents and Licensing Dept
/ STREET: 200 E Randolph St
/ CITY: Chicago
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60680-0703
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.24
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/095,726
/ FILING DATE: 21-JUL-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/785,566
/ FILING DATE: 30-OCT-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Galloway, No. 5530188-va1 B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 3128567180
/ TELEFAX: 3128564972
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 374 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
US-08-095-726-14

```

```

Query Match      35.7%; Score 713.5; DB 1; Length 374;
Best Local Similarity 41.5%; Pred. No. 9.1e-70;
Matches 158; Conservative 55; Mismatches 139; Indels 29; Gaps 6;

QY 4 DLIAGAGLIALAVDRRDPARIWMLDARSGSDQHTWSCDPTLSPMLARLSPI 63
DB 3 DLIAGGGLANGLIARLRQRYPQLMLLIEAGEQPGNHTWSFHEDDITPGQHWLADL 62
QY 64 RRGWTDQVAPDPHSRLITGSGIEAGLIGLQ--GVDLRWNTHTVATLDDTGATLT 120
DB 63 VAHAMPGEVQPPDLRRRLARGYSITSERFAEALHQAALGEMINWNCSEVLPSVRLA 122
QY 121 DGSRIEACVIDARCAVETPHLTGQKFGVGEIETDAPHGVERPMINDATVPQMDGYRF 180
DB 123 NBEALLAGAVIDGRGTATSSAMQGYQLFGQOMRLTQPHGLTVPLMDATVAQOQGYRF 182
QY 181 IYLPFSPTRLIETDRYSDDGDLDDGALAQAASLDYAARRGW-TQGEWRRERGILPIALA 239
DB 183 VYTLPLASADTLIEDTRYANVPQRDNDALRQVTVTDVAHSGMQLAQLREBETGCLPIT-- 240
QY 240 HNAIGWRHQAQAVV-----GLGAGLFHPTGTSIYYAANVAIAA-----ARDLT 287
DB 241 -----WRTSRLICGWRRAASGMAGLFHPTTGSILPLAVALADADSPRLGVSPLY 294
QY 288 TASARAAYGMAIDRADRDRLFLNRMFLRCGCPRRARYLRLQRFRLPOPLIERFYAGR 347
DB 295 QLTROCAEHRM-----RGFRLNRMFLAGREENRWRVWQRFGLBEPIVERFYAGR 349
QY 348 LTLADRLRIVTGRPPILPSQA 368
DB 350 LSLFDRARILTGKPPVPLAKS 370

```

```

RESULT 12
US-08-096-623A-14
/ Sequence 14, Application US/08096623A
/ Patent No. 5684238
/ GENERAL INFORMATION:

```

```

/ APPLICANT: Ausich, Rodney L.
/ APPLICANT: Brinkhaus, Friedhelm L.
/ APPLICANT: Mukharji, Indrani
/ APPLICANT: Proffitt, John H.
/ APPLICANT: Yarger, James G.
/ APPLICANT: Yen, Hwei-Che B.
/ TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
/ TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts
/ NUMBER OF SEQUENCES: 104
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Welsh & Katz, Ltd.
/ STREET: 120 S. Riverside Plaza, 22nd Floor
/ CITY: Chicago
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60606
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/096,623A
/ FILING DATE: 22-JUL-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/805,061
/ FILING DATE: 09-DEC-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/662,921
/ FILING DATE: 28-FEB-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/562,674
/ FILING DATE: 03-AUG-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/525,551
/ FILING DATE: 18-MAY-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/487,613
/ FILING DATE: 02-MAR-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Gamsen, Edward P.
/ REGISTRATION NUMBER: 29,381
/ REFERENCE/DOCKET NUMBER: AMO-006.1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 655-1500
/ TELEFAX: (312) 655-1501
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 374 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
US-08-096-623A-14

```

```

Query Match      35.7%; Score 713.5; DB 1; Length 374;
Best Local Similarity 41.5%; Pred. No. 9.1e-70;
Matches 158; Conservative 55; Mismatches 139; Indels 29; Gaps 6;

QY 4 DLIAGAGLIALAVDRRDPARIWMLDARSGSDQHTWSCDPTLSPMLARLSPI 63
DB 3 DLIAGGGLANGLIARLRQRYPQLMLLIEAGEQPGNHTWSFHEDDITPGQHWLADL 62
QY 64 RRGWTDQVAPDPHSRLITGSGIEAGLIGLQ--GVDLRWNTHTVATLDDTGATLT 120
DB 63 VAHAMPGEVQPPDLRRRLARGYSITSERFAEALHQAALGEMINWNCSEVLPSVRLA 122
QY 121 DGSRIEACVIDARCAVETPHLTGQKFGVGEIETDAPHGVERPMINDATVPQMDGYRF 180
DB 123 NBEALLAGAVIDGRGTATSSAMQGYQLFGQOMRLTQPHGLTVPLMDATVAQOQGYRF 182
QY 181 IYLPFSPTRLIETDRYSDDGDLDDGALAQAASLDYAARRGW-TQGEWRRERGILPIALA 239

```

```
Db      183 VTTLLSADTLLIEBTRYANVPQORDNARQTVTVDYASKQWLAQLEEEETGCLPIT-- 240
      240 HDAIFGWDHAGAVP-----GAGAGFHPVTGSLPYAQNVDATA-----ARDJT 287
      241 -----WRTSYLSCGMRRAASGMAGLFHPTTGLSLPALVALADALDSRLLSVPILY 294
      288 TASARAVAGWAMIDRADRDRLNRLMFRGCPDPRRYRLQRFYRLPOPLIERFYAGR 347
      295 QLTQFPAERHMR-----RQGFRLNRLMFLAGREENRMRVWQRFYGLPEPTVERFYAGR 349
      348 LTLADRLRIVTGRPPILPSQA 368
      350 LSLFDKARILTGKPPVPLAKS 370

RESULT 13
US-09-323-998E-55
; Sequence 55, Application US/09323998E
; Patent No. 6642021
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998E
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 55
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-323-998E-55

Query Match      9.8%; Score 195.5; DB 4; Length 501;
Best Local Similarity 23.4%; Pred. No. 1.6e-12;
Matches 111; Conservative 66; Mismatches 144; Indels 153; Gaps 27;

      4 DLLIAGAGLALALAVDRRPPARIYVLDARSGSPDHTWSCDHTLSPFWLARLSP 63
      84 DLAVGGGAGLAQAQVSE--AGLSVCSID--PSPKLIW-----PN 121
      64 RRGWTDQ-----EVAFPDH--SRRLTTGYSIEAGALIG-----LLQ 99
      122 NVGVWVDFEAMDLLDCLDTTWSGAVVYVDEGVKDLSPRYGRVNRKQKSKMOKCITN 181
      100 GVDLRNN--THVATLDDTQATL--TDGSRIEAACVVDANG-----AVETPHLTGFGQFV 149
      182 GKRFQSKVTNVVH--EENASTVVCSDGVKIQASVLDATGFSRCLVQYQKPY--NPGYQVYA 239
      150 VGVVEIETJA--PHGVERPMIDATVPQMDGY-----RFIYLLPSPFRIILIEDTR 197
      240 YGVVAEVDGHPDQVDMFMWRDKHLDSYPELKERNSKIPFLYIMPPSSNRIFLEET-- 298
      198 YSDGDDLDGALAQAASL-----DYAARRGWGTQGMER--ERGILPILALHDAIGFW 246
      299 -----SLVARPGLRMEDIQERMAARLKHGINVAKRIEEDRCVIMG----- 340
      247 RDHAGAVP-----VGLG--AGLFHPVTGY--SLPYAAQVADAIAARDLTTTSARARA 294
      341 -----GPLFVLPRQVVVGIGTAGWVHPSTGYWVARLAAPIVANAII--RRYIGSSSS--NS 393
      295 VNG-----NAIDRADRFLRLNRLMFRGCPDPRRYRLQRFYRLPOPLIERFYAG 343
      394 LRGDOLSAEAVWMDLPIERRRQREFFCFGMIDLK--LDIDATRRRFFDAFDLQPHYWGCF 452
      344 YAGRLTLAD-----RLRIVTGRPPILPSQAQVAVRCLPERPLQER 381
```

```
Db      453 LSSRLFELLVFGLSLFSHASNTSRLEIMT--KGVPLAKMI-----NNLVQDR 500

RESULT 14
US-09-323-998E-57
; Sequence 57, Application US/09323998E
; Patent No. 6642021
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998E
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 57
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Capsicum annuum
US-09-323-998E-57

Query Match      8.5%; Score 170.5; DB 4; Length 498;
Best Local Similarity 21.4%; Pred. No. 9e-10;
Matches 101; Conservative 67; Mismatches 154; Indels 149; Gaps 26;

      4 DLLIAGAGLALALAVDRRPPARIYVLDARSGSPDHTW-----S 46
      83 DLAVGGGAGLAQAQVSEAGLSVCSIDPPKLIW-----PNNYGVWVDFEAMDLLD 136
      47 CHDITLSPFWLARLSPIRGEWTDQEVAPDHSRR--LTTGYGSIIEAGALIG-----LLQ 99
      137 CLD-----ATWSGAAYVIDKXTQDNLNRPYGRVNRKQKSKMOKCITN 180
      100 GVDLRNNTHVATLDDTQATL--TDGSRIEAACVVDANG-----AVETPHLTGFGQFV 150
      181 GKRFQAVKIVKIHESKSMICNDGITIQATVLDATGFSRSLVQYQKPY--NPGYQVYA 239
      151 GV--EIEDTDPHGVGRPMID-----ATVPQMDGYRFIYLLPSPFRIIL 192
      240 GILAEVE--EHPDVKVPMFMWRDROSHLKNVIELKERNRIP-----TLYMPPSSNRIF 293
      193 IEDTRY--SDGDDLDGALAQAASLDYARRGWGTQGMER--ERGILPILALHDAIGFW 246
      294 LEETSLVARPGLGMD--IQERM--VARLSHLGIKVKSIEDEHCVIIMG----- 339
      247 RDHAGAVP-----VGLG--AGLFHPVTGSLPYAAQVADAIAARDLTTTSARAVNG- 297
      340 -----GPLFVLPRQVVVGIGTAGWVHPSTGYWVARLAAPIVANAIIQYLSSESSHSGD 394
      298 -----NAID--RAARDRFLRLNRLMFRGCPDPRRYRLQRFYRLPOPLIERFYAG 346
      395 ELISAAMVDLMPFIERRRQREFFCFGMIDLKLDLPATRRF--FLAFPLDERYMHGFLSS 452
      347 RLTLAD-----RLRIVTGRPPILPSQAQVAVRCLPERPLQER 381
      453 RLFLBELIVFGLSLFSHASNTSRLEIMT--KGTLLPVHMI-----NNLLQDK 497

RESULT 15
US-09-323-998E-56
; Sequence 56, Application US/09323998E
; Patent No. 6642021
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
```

FILE REFERENCE: 108172-09019
CURRENT APPLICATION NUMBER: US/09/323,998E
CURRENT FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: 09/088,724
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 09/088,725
PRIOR FILING DATE: 1998-06-02
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 56
LENGTH: 502
TYPE: PRF
ORGANISM: Adonis palaeestina
US-09-323-998E-56

Query Match 8.2%; Score 163.5; DB 4; Length 502;
Best Local Similarity 21.1%; Pred. No. 5,4e-09;
Matches 92; Conservative 61; Mismatches 141; Indels 141; Gaps 22;

QY 4 DLLIAGAGLSGALIA-----LAVDRRPDARIWLDARSGPSDQHTW-----S 46
DB 87 DIAVVGCGPAGLAIAGQVSEAGLVCSIDPSPKLIW-----PNNYGVWVDFEAMDLID 140
QY 47 CHDTLSPFWLARIPIRGEWTDQEVAFPPDSRR-LTTGYGIEAGALIG-LIQ----- 99
DB 141 CUDT-----TWGAVVYDDNSKYLDRPYGRVNRKQLSKMLQKCTN 184
QY 100 GVDLRMNTVATLIDTGTATL---TDGSRIEACVIDARG-----AVETPHLTVGFQKFV 150
DB 185 GVKFHQAKIVKIHEESKSLICNDGITINATVLDATGFSRCLVQYDKPY-NPGYQVAY 243
QY 151 GVEIETDA-PHGEVPMIND-----ATVPQMDGYRFTYLLPSPTRILI 193
DB 244 GIMAEVEBHPFDLQKLFMDWRDShLNEKLEKDKNRKIP-----TFIYAMPFSSSTKIPL 298
QY 194 EDTRYSDGDLDDGALAQASLDY-----AARGWTGOEMR---ERGITLPIALAHDA 242
DB 299 EET-----SLVAPGLRPEDIQERWVARLKHIGIKVKSIEEDRCVAPMG----- 343
QY 243 IGFWRDHAQAVP-----VGLG--AGLFHPVYTGYSLPYAAQVADAIARDLTTASARRA 294
DB 344 -----GFLPVLPGQVVGIGGTAGVHPSTGYVAVARTLAAPVAVAKSIVQYLGSDRS 394
QY 295 VAG-----WAIDRADRDF---LRLNMLFRGCPPDRRYRLQRFYRLPQPL 339
DB 395 LSGNELSAEVMKDLWPIERRRQREFFPCFMDILILKLDLG-----TRRFDAFFDLRPHY 449
QY 340 IERFYAGRLTLADRL 354
DB 450 WHGFLSSRLFLPELL 464

Search completed: June 17, 2005, 19:39:10
Job time : 39.3429 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2005, 18:53:39 ; Search time 140.801 Seconds
(without alignments)
1049.298 Million cell updates/sec

Title: US-10-695-980-5

Sequence: 1 MSHDILLIAGAGLGGALIALA.....IPLSQAVRCLPRLPLQERA 382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

A_Geneseq.16Dec04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	382	2	AAW06517
2	1999	100.0	382	2	AAW05953
3	1456	72.8	386	2	AAW79062
4	1456	72.8	386	2	AAW87884
5	1405	70.3	386	2	AAW99096
6	1405	70.3	386	2	ADQ96834
7	810.5	40.5	380	8	ADP74123
8	810.5	40.5	382	2	AAW74465
9	810.5	40.5	382	2	AAW82258
10	810.5	40.5	382	2	AAW87891
11	810.5	40.5	382	2	AAW91100
12	783.5	39.2	382	5	AAE22313
13	783.5	39.2	382	6	AAO16019
14	783.5	39.2	382	6	ABP96687
15	783.5	39.2	382	8	ADQ14629
16	783.5	39.2	382	8	ADQ77204
17	783.5	39.2	382	8	ADQ48639
18	783.5	39.2	382	8	ADQ94456
19	713.5	35.7	374	2	AAW01125
20	713.5	35.7	374	2	AAW24474
21	712.5	35.6	374	2	AAW13987
22	671	27.6	434	6	ABW70122
23	551.5	9.8	501	3	AAW95698
24	195.5	9.8	501	3	AAW54311
25	195.5	9.8	501	3	AAW53411

26	195.5	9.8	501	5	ABB92263	Abp92263 Herbicide
27	194.5	9.7	500	8	ADO05148	Ado05148 A. thalia
28	194.5	9.7	501	8	ADO05224	Ado05224 A. thalia
29	178.5	8.9	411	8	ADH39777	Adh39777 Streptomycin
30	175	8.8	328	3	AAW53413	AAW53413 Arabidopsis
31	175	8.8	369	3	AAW53412	AAW53412 Arabidopsis
32	174.5	8.7	524	8	ADO05152	Ado05152 Citrus X
33	172.5	8.6	524	8	ADO05152	Ado05152 Citrus X
34	170.5	8.5	498	2	AAW06454	AAW06454 Capsicum
35	170.5	8.5	498	3	AAW54313	AAW54313 Amino acid
36	163.5	8.2	502	3	AAW54312	AAW54312 Amino acid
37	161.5	8.1	500	3	AAW54314	AAW54314 Amino acid
38	161.5	8.1	500	8	ADQ61088	Adq61088 Tomato be
39	161.5	8.1	500	8	ADP74116	Adp74116 Tomato be
40	161.5	8.1	500	8	ADQ38257	Adq38257 Tomato be
41	161.5	8.1	500	8	ADQ94526	Adq94526 Tomato be
42	158	7.9	498	3	AAW70397	AAW70397 Protein e
43	157.5	7.9	511	3	AAW90226	AAW90226 Marigold
44	157.5	7.9	511	3	AAW54316	AAW54316 Amino acid
45	157	7.9	500	3	AAW54315	AAW54315 Amino acid

ALIGNMENTS

RESULT 1

AAW06517 standard, protein, 382 AA.

XX	AAW06517;				
AC	17-OCT-2003	(revised)			
DT	08-MAR-1997	(first entry)			
DE	Flavobacterium lycopen cyclase.				
KW	Carotenoid; lycopene; beta-carotene; echinenone; canthaxanthin;				
KW	zeaxanthin; adonixanthin; astaxanthin; crt; lycopene synthase.				
OS	Flavobacterium sp. ATCC 21588; WT (ATCC 21588).				
PN	EP747483-A2.				
XX	11-DEC-1996.				
PD	29-MAY-1996;	96EP-00108556.			
PF	09-JUN-1995;	95EP-00108888.			
PR	(HOF) HOFFMANN LA ROCHE & CO AG F.				
PA	Hohmann H, Pasamontes L, Tessier M, Van Loon A;				
PI	WPI; 1997-023160/03.				
DR	N-PSDB; AAT45143.				
XX	Flavobacterium gene sequences encoding carotenoid biosynthesis enzymes -				
PT	for the production of carotenoid(s), useful in foods and animal feeds.				
XX	Example 2; Fig 7; 80pp; English.				
XX	Lycopene synthase (AAW06517) is the product of the crt gene identified				
CC	in a carotenoid gene cluster (see also AAT45143) of Flavobacterium sp.				
CC	R1534. This 42368 Da protein introduces the beta-ionone rings at both				
CC	sides of lycopene to obtain beta-carotene. Enzymes of the Flavobacterium				
CC	carotenoid biosynthetic pathway (see also AAW06515-18 and AAW00871) can				
CC	be expressed, optionally with Alcaligenes beta-carotene beta-4-oxygenase,				
CC	in host cells for the prodn. of lycopene, astaxanthin, beta-carotene,				
CC	echinenone, zeaxanthin, canthaxanthin and adonixanthin. (Updated on 17-				
CC	OCT-2003 to standardise OS field)				
XX	Sequence 382 AA;				
SQ					

Query Match 100.0%; Score 1999; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 9.6e-201;
 Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLIAGAGISGALIALAVDRRPPDARIYMLDARSGPSDOHTWSCDITLSPFWLRL 60
 DB 1 MSHDLIAGAGISGALIALAVDRRPPDARIYMLDARSGPSDOHTWSCDITLSPFWLRL 60

QY 61 SPIRGEMTDOEAVAPDHSRLTTGYSIEAGALIGLQGVLDLRNTHVATLDDTGATLT 120
 DB 61 SPIRGEMTDOEAVAPDHSRLTTGYSIEAGALIGLQGVLDLRNTHVATLDDTGATLT 120

QY 121 DGSRIEACVIDARCAVETPHLTGFGKFGVGEIETDAPHGVERPMINDATVPQMDGYRF 180
 DB 121 DGSRIEACVIDARCAVETPHLTGFGKFGVGEIETDAPHGVERPMINDATVPQMDGYRF 180

QY 181 IYLLPSPFTRILLIEDTRYSDGDLDDGALAQASLDYAARRGWTGQEMRREGILPILALH 240
 DB 181 IYLLPSPFTRILLIEDTRYSDGDLDDGALAQASLDYAARRGWTGQEMRREGILPILALH 240

QY 241 DAIGFWRDHAQAVPVGLGAGLFHPVTGYSLPYAQVADAIARDLTTASARRAVGMAI 300
 DB 241 DAIGFWRDHAQAVPVGLGAGLFHPVTGYSLPYAQVADAIARDLTTASARRAVGMAI 300

QY 301 DRADRDRLRLINRMLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360
 DB 301 DRADRDRLRLINRMLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360

QY 361 PPILPSQAVRCLPERPILQERA 382
 DB 361 PPILPSQAVRCLPERPILQERA 382

RESULT 2
 AAW69533 standard; protein; 382 AA.
 XX AAW69533;
 AC XX
 XX 10-AUG-1999 (first entry)
 DT XX
 DE Flavobacterium sp. R1534 crtY gene product lycopen cyclase.
 XX
 KW Carotenoid; pigment; canthaxanthin; R1534; crtB; prephytoene synthase;
 KM crtI; phytoene desaturase; crtY; lycopene cyclase; GGPP synthase; crtE;
 XX crtW2396; beta-carotene beta-oxygenase; food product; fermentation.
 OS Flavobacterium sp.
 XX
 PN JP10155497-A.
 XX
 PD 16-JUN-1998.
 XX
 PF 02-DEC-1997; 97JP-00348653.
 XX
 PR 02-DEC-1996; 96BP-00810839.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 DR WPI; 1998-391048/34.
 DR N-PSDB; AAV40146.
 XX
 PT Preparation of carotenoid - comprises fermentation with transformed cell.
 XX
 PS Claim 1; Fig 25; 80pp; Japanese.
 CC The invention describes the preparation of carotenoid pigments e.g.
 CC canthaxanthin using a cell transformed by a vector having DNA sequences
 CC (a) to (e) or substantially homologous sequences. (a) a DNA sequence
 CC (crtE) coding GGPP synthase of Flavobacterium sp. R1534; (b) a DNA
 CC sequence (crtB) coding prephytoene synthase of Flavobacterium sp. R1534;
 CC (c) a DNA sequence (crtI) coding phytoene desaturase of Flavobacterium
 CC sp. R1534; (d) a DNA sequence (crtY) coding lycopene cyclase of

CC Flavobacterium sp. R1534, and (e) a DNA sequence (crtW2396) coding beta-
 CC carotene beta-oxygenase of a microbe E-396 (FERM BP-4283). The carotenoid
 CC or a carotenoid mixture can also be used in preparation of food products.
 CC The method is an improved method of fermentation for carotenoid
 CC production
 XX
 SQ Sequence 382 AA;

Query Match 100.0%; Score 1999; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 9.6e-201;
 Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHDLIAGAGISGALIALAVDRRPPDARIYMLDARSGPSDOHTWSCDITLSPFWLRL 60
 DB 1 MSHDLIAGAGISGALIALAVDRRPPDARIYMLDARSGPSDOHTWSCDITLSPFWLRL 60

QY 61 SPIRGEMTDOEAVAPDHSRLTTGYSIEAGALIGLQGVLDLRNTHVATLDDTGATLT 120
 DB 61 SPIRGEMTDOEAVAPDHSRLTTGYSIEAGALIGLQGVLDLRNTHVATLDDTGATLT 120

QY 121 DGSRIEACVIDARCAVETPHLTGFGKFGVGEIETDAPHGVERPMINDATVPQMDGYRF 180
 DB 121 DGSRIEACVIDARCAVETPHLTGFGKFGVGEIETDAPHGVERPMINDATVPQMDGYRF 180

QY 181 IYLLPSPFTRILLIEDTRYSDGDLDDGALAQASLDYAARRGWTGQEMRREGILPILALH 240
 DB 181 IYLLPSPFTRILLIEDTRYSDGDLDDGALAQASLDYAARRGWTGQEMRREGILPILALH 240

QY 241 DAIGFWRDHAQAVPVGLGAGLFHPVTGYSLPYAQVADAIARDLTTASARRAVGMAI 300
 DB 241 DAIGFWRDHAQAVPVGLGAGLFHPVTGYSLPYAQVADAIARDLTTASARRAVGMAI 300

QY 301 DRADRDRLRLINRMLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360
 DB 301 DRADRDRLRLINRMLFRGCPDPRRYRLQRFYRLPQPLIERFYAGRLTLADRLRIVTGR 360

QY 361 PPILPSQAVRCLPERPILQERA 382
 DB 361 PPILPSQAVRCLPERPILQERA 382

RESULT 3
 AAR79062 standard; peptide; 386 AA.
 ID AAR79062
 AC AAR79062;
 XX
 DT 27-AUG-2003 (revised)
 DT 28-FEB-1996 (first entry)
 XX
 DE 3 hydroxy-beta-ionone ring methylene to keto group converting peptide.
 XX
 KW Xanthophyll; astaxanthine; methylene; keto group; conversion;
 KM 3-hydroxy-beta-ionone ring.
 XX
 OS Agrobacterium aurantiacum.
 XX
 PN WO9518220-A1.
 XX
 PD 06-JUL-1995.
 XX
 PF 26-DEC-1994; 94WO-JP002220.
 XX
 PR 27-DEC-1993; 93JP-00348737.
 XX
 PR 05-SEP-1994; 94JP-00235917.
 XX
 PA (KIRI) KIRIN BEER KK.
 PA (MARI-) MARINE BIOTECHNOLOGY INST CO LTD.
 XX
 PI Misawa N, Kondo K, Kajiwara S, Yokoyama A;
 XX
 DR WPI; 1995-246386/32.
 DR N-PSDB; AAG99489.

XX DNA's encoding xanthophyll(s) - esp. asta:xanthin and other
 PT xanthophyll(s) using e.g. E. coli.
 XX
 PS Disclosure, Page 21; 131pp; Japanese.
 XX
 CC AAR79058-R790629 are xanthophyll polypeptides. These polypeptides are
 CC capable of converting the 4-methylene group of a 3-hydroxy-beta- ionone
 CC ring to a 4-keto group in doing so these peptides also add a hydroxyl
 CC group to the 3-position carbon-atom of the 4-keto-beta- ionone ring. The
 CC DNA sequences may be used in the production of astaxanthin and other
 CC keto gp. contg. xanthophylls, the sequences may also be used to transform
 CC certain yeasts and other microorganisms. (Updated on 27-AUG-2003 to
 CC correct OS field.)
 CC
 XX Sequence 386 AA;
 SQ
 Query Match 72.8%; Score 1456; DB 2; Length 386;
 Best Local Similarity 71.8%; Pred. No. 1.1e-143;
 Matches 277; Conservative 39; Mismatches 66; Indels 4; Gaps 2;
 QY 1 MSHDLLIAGAGLSGALITAAVRRPDAIWMIDARSGSDQHTWCHDITLSPWLAAL 60
 DB 1 MTHDVLIAAGAGLANGLIALALRAAPDLRLVLLDHAAGSDGHTWCHDITLSPWLAAL 60
 QY 61 SPIRGEMTDQEVAPPDHSRRLTTGYGSTEAGALIGLL--QGVDLRMTHTVATLDDTGAT 118
 DB 61 KPLRRANMPDQEVRRPRAARLRTATGSLDGAALDAVVRSGAEIRMSDIALLDAGAT 120
 QY 119 LFDGSRIEACVTDARGAVERPHLTGFOKFGVEIETDAPHGVERPMIMDATVPMQDGY 178
 DB 121 LSCGTRIEAGAVLDGGAQPSRLTVGFOKFGVEIETDRPHGVERPMIMDATVPMQDGY 180
 QY 179 RFTYLLPFSPTRLIEDTRYSDGDLDDGALAQASLDYARRGWTQEMRRERGIPIAL 238
 DB 181 RFTYLLPFSPTRLIEDTRYSDGDLDDGALAASHDYARRGWTGAVERRRERGIPIAL 240
 QY 239 AHDAIGFMDHQAQAVPGAGLPHPTGYSLPYAAQVADVA--ARDLTTASARRAVR 296
 DB 241 AHDAIGFMDHQAQAVPGAGLPHPTGYSLPYAAQVADVAAGSGPGTDALRGAR 300
 QY 297 GNAIDRADRDRLRLNRLFRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLADRLRI 356
 DB 301 DYALDARDRRLRLNRLFRGCAPDRRYTLQRFYRPHGLIERFYAGRLSVADQLRI 360
 QY 357 VTGRPPILSQAVRCLPERPLQERA 382
 DB 361 VTGKPPILPGTAIRCLPERPLKENA 386
 RESULT 4
 AAW87884
 ID AAW87884 standard; protein; 386 AA.
 XX
 AC AAW87884;
 XX
 DT 10-MAR-1999 (first entry)
 XX
 DE Protein encoded by the carotenoid biosynthesis gene crtY.
 XX
 KW Carotenoid biosynthesis; astaxanthin diglucoside; crtE gene; crtB gene;
 KW crtI gene; adonixanthin-3'-glucoside; astaxanthin monoglucoside;
 KW carotenoid glycoside; crtY gene; crtZ gene; crtX gene; crtW gene;
 KW food additive.
 XX
 OS Agrobacterium aurantiacum.
 XX
 PN JP10327865-A.
 XX
 PD 15-DEC-1998.
 XX
 PF 29-MAY-1997; 97JP-00140460.
 XX

PR 29-MAY-1997; 97JP-00140460.
 XX
 PA (KIRI) KIRIN BREWERY KK.
 PA (KAIT-) KAITO BIOTECHNOLOGY KENKUTSHO KK.
 XX
 DR WPI; 1999-099030/09.
 DR N-PSDB; AAW87884.
 PT New carotenoid glucoside(s) - used as food additives.
 XX
 PS Disclosure, Page 12-13; 26pp; Japanese.
 XX
 CC The present sequence represents a protein involved in carotenoid
 CC biosynthesis. The specification describes astaxanthin diglucosides and
 CC adonixanthin-3'-glucosides. The specification also describes a method for
 CC the preparation of a carotenoid glycoside, in which all, or part of,
 CC carotenoid biosynthesis genes crtE, crtB, crtI, crtY, crtZ, crtX or crtW
 CC are introduced to a microbe or plant and expressed. The transformed
 CC organism is cultured and astaxanthin diglucosides, adonixanthin-3'-
 CC glucosides, and/or astaxanthin monoglucosides are collected. The
 CC carotenoid glucosides are used as food additives
 CC
 XX Sequence 386 AA;
 SQ
 Query Match 72.8%; Score 1456; DB 2; Length 386;
 Best Local Similarity 71.8%; Pred. No. 1.1e-143;
 Matches 277; Conservative 39; Mismatches 66; Indels 4; Gaps 2;
 QY 1 MSHDLLIAGAGLSGALITAAVRRPDAIWMIDARSGSDQHTWCHDITLSPWLAAL 60
 DB 1 MTHDVLIAAGAGLANGLIALALRAAPDLRLVLLDHAAGSDGHTWCHDITLSPWLAAL 60
 QY 61 SPIRGEMTDQEVAPPDHSRRLTTGYGSTEAGALIGLL--QGVDLRMTHTVATLDDTGAT 118
 DB 61 KPLRRANMPDQEVRRPRAARLRTATGSLDGAALDAVVRSGAEIRMSDIALLDAGAT 120
 QY 119 LFDGSRIEACVTDARGAVERPHLTGFOKFGVEIETDAPHGVERPMIMDATVPMQDGY 178
 DB 121 LSCGTRIEAGAVLDGGAQPSRLTVGFOKFGVEIETDRPHGVERPMIMDATVPMQDGY 180
 QY 179 RFTYLLPFSPTRLIEDTRYSDGDLDDGALAQASLDYARRGWTQEMRRERGIPIAL 238
 DB 181 RFTYLLPFSPTRLIEDTRYSDGDLDDGALAASHDYARRGWTGAVERRRERGIPIAL 240
 QY 239 AHDAIGFMDHQAQAVPGAGLPHPTGYSLPYAAQVADVA--ARDLTTASARRAVR 296
 DB 241 AHDAIGFMDHQAQAVPGAGLPHPTGYSLPYAAQVADVAAGSGPGTDALRGAR 300
 QY 297 GNAIDRADRDRLRLNRLFRGCPDRRYRLQRFYRLPOPLIERFYAGRLTLADRLRI 356
 DB 301 DYALDARDRRLRLNRLFRGCAPDRRYTLQRFYRPHGLIERFYAGRLSVADQLRI 360
 QY 357 VTGRPPILSQAVRCLPERPLQERA 382
 DB 361 VTGKPPILPGTAIRCLPERPLKENA 386
 RESULT 5
 AAW99096
 ID AAW99096 standard; protein; 386 AA.
 XX
 AC AAW99096;
 XX
 DT 14-MAY-1999 (first entry)
 XX
 DE Agrobacterium aurantiacum crtY protein sequence.
 XX
 KW Beta-carotene hydroxylase; crtY; crtB; crtE; crtI; xanthophyll;
 KW metabolite.
 XX
 OS Agrobacterium aurantiacum.
 XX
 PN JP1046770-A.
 XX

DE Pantoea ananatis CrtY.
XX
XX
KW ketocarotenoid; ketolase; carotenoid; hydroxylase; beta-cyclase;
XX astaxanthin; adonirubin; canthaxanthin; adoxanthin; zeaxanthin;
KW antioxidant; pigment; animal feed; trout; salmon; shrimp;
XX food supplement; enzyme.
XX
OS Pantoea ananatis.
XX
XX DE10253112-A1.
XX
XX 03-JUN-2004.
XX
XX 13-NOV-2002; 2002DE-01053112.
XX
XX 13-NOV-2002; 2002DE-01053112.
XX
XX (SUNG-) SUNGENE GMBH & CO KGAA.
XX
XX Sauer M, Flachmann R, Klebatschel M, Schopfer CR;
XX
XX WPI; 2004-469014/47.
XX
XX N-PSDB; ADP74122.
XX
XX Production of ketocarotenoids with low hydroxylated by-product content,
XX for use e.g. in pigmenting feedstuffs, by culturing genetically modified
XX organisms having modified ketolase activity.
XX
XX Example 2; SEQ ID NO 25; 101pp; German.
XX
XX This invention describes a novel method for the production of
XX ketocarotenoids which involves culturing genetically modified organisms
XX having modified ketolase (KLA) activity. The starting microorganisms
XX produce carotenoid (naturally or by genetic supplementation), and are
XX specifically microorganisms (especially bacteria, yeasts, algae or fungi)
XX or plants. Microorganisms and plants involved in the method could include
XX Escherichia, Flavobacterium, Nostoc, Synochocystis, Haemulna, Fusarium,
XX Dunaliella, Rannunciacae, Cannabaceae, Brassicaceae, Amaranthaceae,
XX Solanaceae, Lamiaceae, Acacia, Calendula, Gentiana, Helianthus, Linum,
XX Rhododendron, Spartium and Zinnia. The modified microorganisms
XX additionally show elevated hydroxylase and/or beta-cyclase activity. The
XX ketocarotenoids described in the invention include astaxanthin,
XX adonirubin, canthaxanthin, adoxanthin and zeaxanthin. Ketocarotenoids are
XX natural antioxidants and pigments, especially useful as pigmenting
XX additives in animal feed, specifically feed for trout, salmon or shrimps.
XX Ketocarotenoid-producing genetically modified organisms can be used as
XX feedstuffs or foodstuffs, in the production of ketocarotenoid-containing
XX extracts or for producing feed or food supplements. The process provides
XX large amounts of ketocarotenoid having a low content of hydroxylated by-
XX products, especially in the case of astaxanthin, preferably due to
XX expression of at least one nucleic acid encoding hydroxylase and/or beta-
XX cyclase. DNA encoding the whole primary ketolase sequence from Nostoc sp.
XX strain PCC7120 was isolated, amplified by PCR and used to produce a
XX plasmid pNOSTP-G. A plasmid pMCL-Crt-YIBZ/idi/gps, for the synthesis of
XX zeaxanthin in Escherichia coli, was constructed in 3 stages via the
XX intermediate stages pMCL-CrtYIBZ and pMCL-CrtYIBX/idi, using the high
XX copy number plasmid vector pMC200. Escherichia coli strain TOP10 was
XX transformed with the plasmide pNOSTP-G and pMCL-Crt-YIBZ/idi/gps to give
XX carotenoid producing strain.
XX
XX Sequence 380 AA;

Query Match	40.5%	Score 810.5	DB 8	Length 380
Best Local Similarity	45.6%	Pred. No. 6.5e-76		
Matches 171	Conservative 59	Mismatches 138	Indels 7	Gaps 4

[illegible]

Db 63 LVVHHMDYQVAFPIRRRLKNSGVCITISQFAEVLQKQFPHMLMDTAAVAAVBAEVL 124

Qy 120 TDGSRLEAACVLDARGAVETPHLTYGQKFGVGELETDARGVERPMIMATTPQMDNGYR 179

Db 125 KKGQVGAAGAVDGRGYAANSALSVGFQAFIGQERLSSHPLGSLSPIMATVADQOQNGYR 184

Qy 180 FYILPFPSPTRLIEDTRYSDGDDLDALAQASLDYAAARRGWTGQGE-MRERERITLPIAL 238

Db 185 FVYSLPFLSPTRLIEDTTHIDNALTDPECAQNICDYAAQGGWLQTLRLREGQALPITL 244

Qy 239 AHDAIGFMDHDAQAVPGCAGLEPHPTGVSLEPAAQVAAALAAARDL-TAASARAARVG 297

Db 245 SGNAADFWQORPLAC--SGLRAGLPHPTGVSLEPAAVAAARLALDLDVFASISIHNAITH 302

Qy 298 WAIDRADRDRLFLILNRMLFRCCPPDRRYRLQRFYRLPOPLIERFYAGRLTLADRLRIY 357

Db 303 FAERHQOQGFPRMLNRMLFLAGPADSEWRVMQRFYGEEDLIARFYAGKLLTDLRLIL 362

Qy 358 TGRPPILSLQAVRCL 372

Db 363 SGKEPVFVLAALQAI 377

RESULT 8
AAR07465
ID AAR07465 standard; protein; 382 AA

AC AAR07465

DT	24-OCT-2003	(revised)
DT	28-JAN-1991	(first entry)

DE Polypeptide with enzymatic activity for the conversion of lycopene into
DE beta-carotene.
DE

KW Carotenoid biosynthesis; vitamin A; cancer; food coloring

OS *Pantoea ananatis*.

PN EP393690-

PD 24-OCT-1990

PF 20-APR-1990; 90EP-00107493.

PR	21-APR-1989;	89JP-00103078.
PR	05-MAR-1990;	90JD-00053355

XX
PA (YBT \ YBTN DEEB WY

XX	
PI	Misawa N Kobayashi K Nakamura K.

XX
DR
WPT: 1990-322212/43

DK N-PSDB; AAQ06295
XX

PT of carotenoid cpds, e.g. beta-carotene, lycopene, phytoene, etc.

Claim 3: Fid 3: 40pp: English.

Gene products are useful for the synthesis of carotenoids, useful as food

also AAQ06293-9. (Updated on 24-OCT-2003 to standardise OS field)

SQ Sequence 382 AA;

Query Match	40.5%	Score 810.5	DB 2	Length 382
-------------	-------	-------------	------	------------

Matches 171; Conservative 59; Mismatches 138; Indels 7; Gaps 4;

Oy 3 HDLLIAGGLSGALIALAVRDREPDARIYMLDARSPDSQHTWSCIDTDLISPEWLRLSP 62
 ::
Dd 5 YDILIVGGLANGLIALRLQQQRPMRILLIDARPQGSHNTWSFFHDDLTESQHRIAP 64

Qy	65	IRGEWTDGVAAPDHSRLTTGYGSIKAGLIGLQ---	GVDLRMNTHVATLDDTATL	119
Db	65	LVAHHMEDYVRRPPTRRKINSGYFITTSQRPAEVLQRF	GHGLMMDTAALVAEVAESVRL	124
Qy	120	TDSRIEAACTIDARGAVETPHLTVGFOKEVVEI	ETETAPGVERPMTADATVQMDGYR	179
Db	125	KKGCVIGARAVIDGRGAANSALSVQFQFIQGEWR	LSPHPLGSLPIIMDTVTDQNGYR	184
Qy	180	FYLLPESPTRILIEPTRVSDGDDLDGALAAQSL	ADYAAARGMTGOE--MRERGIPLIAL	238
Db	185	FVYSLPESPLRIIEDTHYIDNATLTDPEGARQ	NICDYAAQSGWQQTLLREEGALITL	244
Qy	239	AHDAIGWRDHAQSAVPVIGAGLEFHPVTGSL	PYAAQVADALAAARDI--TTASARRAVRG	297
Db	245	SGNADAPWQQRPLAC--SGLRAGLF	PHPTGYSLPLAVANADRLSLDVFSTASIHNAITH	302
Qy	298	WALDRADRDYFLNLNRMFLFRGCP	PPRRKRLIQRTYRLPOLPIEFYAGRLTLADRIRIV	357
Db	303	FARERWQQQGFEMLNRMFLFLAGPADSR	WRWQRFYGPEDLIARFAGKTLTLDRIRIL	362
Qy	358	TGRPPIPLSQAVRCL	372	
Db	363	SGKRPVEVLAALQAI	377	

RESULT 9
AAW82258
ID AAW82258 standard; protein; 382 AA

ID AA082258 standard; protein; 382 AA.
 AC AA082258;
 XX
 XX 17-OCT-2003 (revised)
 DT 16-JUL-1999 (first entry)
 XX
 XX C. utilis crtY protein.
 DE
 XX
 XX HMG-CoA; 3-hydroxy-3-methylglutaryl coenzyme A reductase; crtY;
 KW carotenoid.
 XX
 XX Pichia jadinii.
 OS
 XX JP10248575-A.
 PN
 XX 22-SEP-1998.
 PD
 XX 12-MAR-1997; 97JP-00058012.
 PF
 XX 12-MAR-1997; 97JP-00058012.
 PR
 XX
 XX (KIRI) KIRIN BREWERY KK.
 PA
 XX WPI; 1998-560727/48.
 DR N-PSDB; AAV73182.
 XX
 XX Gene useful for increase in carotenoid production - and preparation of
 PT carotenoid.
 XX
 XX Example 2; Fig 15-17; 54pp; Japanese.
 PS
 XX This invention describes a novel method for the preparation of
 CC carotenoids using genes and proteins isolated from Candida utilis. The
 CC invention specifically describes the isolation of a 3-hydroxy-3-
 CC methylglutaryl coenzyme A (HMG-CoA) reductase protein. This sequence
 CC represents the Candida utilis crtY protein which is used in the method of
 CC the invention. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 382 AA;
 SQ

Query Match	40.5%	Score	810.5	DB 2	Length	382
Best Local Similarity	45.6%	Pred. No.	6.6e-76			
Matches 171; Conservative	59	Mismatches	138		Indels	7
					Gaps	4

```

Oy      HDLLIAGAGISGLTALAPDRRPARIVMLDASSGPDQHTGSCDTHDSPEWILASP 62
Db      3 YDILILGAGIANGLIALRLQOOQPMRKLILLDAPARQAGSNHTTGSFHHDDLTDESQHRILAR 64
Oy      63 IRGEWTDQEVAPPDHSRRLTTGYGISEAGALIGLQ--GVDLRNTHVATLDDTGATL 119
Db      65 LVVHNHPDVQVAFPTRRKRLNSGYCITISQGFPAEVLQGFQFPHLMQMTAVAAVVAESVRL 124
Oy      120 TDGSRLEAACVIDARGAVETPHLITVGFOKFTVGVEIETDAPRGVERPMIMQATVPQNDGYR 179
Db      125 KKQGVIGAAVAVIDGGAVALNSALISVGQAFIGOEURLSHRPGLSPPIMATAYDOQNGYR 184
Oy      180 FYLLPEPSPRLTIEDTRYSDGGLDQDGAALQASLDVAARGWGTQGE-MRREGILPIAL 238
Db      185 FVYSLPSLSPRLTIEDTRYHIDATLDDPEACARONICTYAAQOGMQLTLLIEBQGAEPITL 244
Oy      239 AHAIGFWEHDQAQAVPVGIGAGLPHPTGYSLPYAAQVADAIALARBL-TTASARAVRG 297
Db      245 SGNAADAFWQORPLAC--SGLRAGLPHPTGYSLPYAAVAVDRLBALDVPFASGIIHNAITH 302
Oy      298 WAIDRADRDRLFLRLINRMLFRGCPDRRYRLLOREFYRLPOPLIRFPYAGRLTLADRIRIV 357
Db      303 FAEERWQOQGFPRMLNRMLFLAGPADSRWRVMORFGYGPBEDLIRAFYAGKTLTDRIRIL 362
Oy      358 TGRPPPIPLSQAVRCL 372
Db      363 SGRPPVPVLAALQAI 377

```

RESULT 10	.
AAW87891	
ID	AAW87891 standard; protein; 382 AA

AC AAW87891;

DT	17-OCT-2003 (revised)
DT	10-MAR-1999 (first entry)
XX	Protein encoded by the carotenoid biosynthesis gene <i>crtY</i>
DE	

KM Carotenoid biosynthesis; astaxanthin diglucoside; crtB gene;
KM crtI gene; adonixanthin-3'-glucoside; astaxanthin monoglucoside;
KM carotenoid glycoside; crtZ gene; crtW gene;
KM food additive.

OS Pantoea ananatis.

PN JP10327865-A.

PD 15-DEC-1998

PF 29-MAY-1997; 97JP-00140460.

PR 29-MAY-1997; 97JP-00140460.

PA (KIRI) KIRIN BREWERY KK.

XX

DR N-PSDB; AAV84083.

PT New carotenoid glucoside(s) - used as food additives.

PS Disclosure; Page 22-23; 26pp; Japanese.

CC The present sequence represents a protein involved in carotenoid
CC biosynthesis. The specification describes astaxanthin diglucosides and
CC adonixanthin-3'-glucosides. The specification also describes a method for
CC the preparation of a carotenoid glycoside, in which all, or part of,
CC carotenoid biosynthesis genes *crte*, *crtb*, *crty*, *crtz*, *crx* or *crw*
CC are introduced to a microbe or plant and expressed. The transformed
CC organism is cultured and astaxanthin diglucosides, adonixanthin-3'-
CC glucosides, and/or astaxanthin monoglucoside are collected. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 19:42:50 ; Search time 4528.7 Seconds

(without alignments)
9657.484 Million cell updates/sec

Title: US-10-695-980-1_COPY_5794_6942

Perfect score: 1149
Sequence: 1 tcatgctctctctcgcagca.....atcagcagatcatgctcat 1149

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hnc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsat:*
9: gb_gsat2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	81.4	7.1	1180	9	CL513526 SAIL_876
C 2	75.4	6.6	1798	9	AG171124 Pan trogl
C 3	75.2	6.5	932	9	CNS00720
C 4	73	6.4	1577	7	CK419311 ANF IPova
C 5	72.6	6.3	1278	9	AG060116 Pan trogl
C 6	71	6.2	1798	9	AG171124 Pan trogl
C 7	70.6	6.1	935	9	CNS006XK
C 8	70	6.1	1821	9	CL090560
C 9	69.8	6.1	1189	9	AG030608 Pan trogl
C 10	69.6	6.1	1406	9	CG756569 P051-4-80
C 11	69	6.0	938	9	AG036223 Pan trogl
C 12	69	6.0	1090	6	CA790876 AGENCOURT
C 13	68.2	5.9	929	9	CL501227 SAIL_694
C 14	68	5.9	896	9	AG159205 Pan trogl
C 15	68	5.9	1250	9	AG043469 Pan trogl
C 16	67.6	5.9	925	9	CNS00919
C 17	67.6	5.9	1057	9	CNS04680
C 18	67.6	5.9	1348	9	CG752544 P047-2-C1
C 19	67.4	5.9	1375	2	AW727483 CA_Ra001
C 20	67.4	5.9	1448	8	CC194731 CH261-183
C 21	67.4	5.9	2193	9	CL469745 SAIL_133
C 22	67.2	5.8	1216	9	AG441574 Mus muscu
C 23	67.2	5.8	1309	9	AG077201 Pan trogl
C 24	66.2	5.8	1244	4	BG846745 1024014HO

25	66.2	5.8	1413	9	CL505027
C 26	66.2	5.8	2332	9	AG363333 Mus muscu
C 27	65.6	5.7	1244	4	BG846745 1024014HO
C 28	65.4	5.7	633	9	AG043206
C 29	65.4	5.7	932	9	CNS00720
C 30	65.2	5.7	1037	9	AG061836
C 31	65.2	5.7	1100	9	CNS01680
C 32	65.2	5.7	1821	9	CL090560
C 33	64.6	5.6	467	9	CG058960
C 34	64.6	5.6	654	8	CC438940
C 35	64.6	5.6	667	9	CG141823
C 36	64.6	5.6	733	8	BZ989278
C 37	64.6	5.6	883	8	CC401856
C 38	64.6	5.6	902	9	CG121335
C 39	64.6	5.6	914	8	CG378162
C 40	64.6	5.6	923	9	CG327267
C 41	64.6	5.6	983	8	CC387908
C 42	64.6	5.6	1341	9	AG030611 Pan trogl
C 43	64.2	5.6	925	9	CNS00919
C 44	64	5.6	1462	9	AG441877 Mus muscu
C 45	64	5.6	1538	9	AG030607 Pan trogl

ALIGNMENTS

RESULT 1
CL513526/c
LOCUS SAIL_876_G03.v2 1180 bp DNA linear GSS 01-APR-2004
DEFINITION SAIL_876_G03.v2, genomic survey sequence.
ACCESSION CL513526
VERSION CL513526
KEYWORDS GT:46010846
SOURCE GSS
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1180)
Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
Bullis,D., Snell,J., Mignel,T., Hutcheson,D., Kimerly,B.,
Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
A high-throughput Arabidopsis reverse genetics system

REFERENCE
AUTHORS Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
Bullis,D., Snell,J., Mignel,T., Hutcheson,D., Kimerly,B.,
Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
TITLE A high-throughput Arabidopsis reverse genetics system
JOURNAL Plant Cell 14 (12), 2985-2994 (2002)
MEDLINE 22356987
PubMed 12468722
COMMENT Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS839448; T-DNA left border flanking sequences of
through the Arabidopsis Insertion library (SAIL) lines are available
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.
FEATURES
source
1..1180
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_876_G03.v2"
/clone.lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a
modified TAIL-PCR strategy"

ORIGIN
Query Match 7.1%; Score 81.4; DB 9; Length 1180;
Best local Similarity 44.9%; Pred. No. 4.2e-07;

	Matches	492: Conserved	0: Mismatches	592: Indels	12: Gaps	5
QY	32	GGCAGCAGCGCAGCGCCTTCGACAGCGGATGCGCGGCGCTCCGGTACGATCGCAAGC				91
Db	1108	GGCGCGCGAGAGGCGAGGCGCGCTGTGTGCGAGTTCGCGCGCGCGCGCGAGATTGCGAGCGG				1045
QY	92	CGGTGCGGCATGTCAAGGCGCCCGGCAATGAAAGCGCTCGATTCAGGGGCTGCGGCAAGCGG				151
Db	1048	GGGGGGGCCCCCAGTGGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG				989
QY	152	TAGAACCCCTGCAGCAGGCGGATAGCAGCGGTGCGGCGGAGCGCAGCGCGGAAACGACTCCGG				211
Db	988	CGGTGCGGCGCGCGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG				929
QY	212	TTTCAGCAGCGCGAGAAAGCGGTTCGCGATTCGCGCGATGATGAGCCGACCGCGCACCGG				271
Db	928	CGAGAGGGGCGCGCGCGGCG				870
QY	272	CGACGCGGCGGAGACGCGGTGTGTAAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG				331
Db	869	CG				810
QY	332	TAGGCGCAGCGAATATTCGGGTGACGGGGGTGGAACACCGCTTCGCCCGCCCAACCGGACAC				391
Db	809	TCGGGCGCACCGGCGG				750
QY	392	GC-----CCCTCTGCGCGTGTGTGTCGCGCGAGAAAGCTATGCGGTTCATGCGGACGCGGATGG				446
Db	749	GCGGGTTCTTCGCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTGT				690
QY	447	GCAGGATGCGCTTTCGCGCGCGCAATCTCTGTGCGCGGTTCAGGCGCGCGCTGTGAGCGCGTAT				506
Db	689	GCGGCG				631
QY	507	CCAGCGACGCGCTGCGCGCAGCGCGCGCGCATGCTTCAGATTCGCGCGCGCGCTGTGAGCGCGTAT				566
Db	630	GCGG				571
QY	567	CCTCGATCAGAGTGCGGGTGGGACTGAAGGCGAGAGATGATGTAAGCGGTACCCGCTCA				626
Db	570	GCGCGGCG				511
QY	627	TCTGCGGAAAGGTGCGCGTTCATGATCATTCGAGCGCTCCAGCGCGCATGGGAGGCGTGGGTCT				686
Db	510	GCGCGGCG				451
QY	687	CGATTCGACGCGCGCAGCAATTTCTGAAAAACCCAGCGTCAAGGTGAGGCGTTCGACGCGCAC				746
Db	450	CGG				394
QY	747	CACGGGCGTTCGATTCACGCGAGCGCGCTCGATTCGCGAGCGCGTCCGTCAGCGTTCGCGCGG				806
Db	393	GCGGGGGGGCGCGCGGCGCGGGCGGGGGGGCGGGGAGACCCCCCGCGGCGCCCCCG				334
QY	807	TATCGTCCAGCGGTGCGCGACATCGGATTTTCACCGCAGATTCAGACATCTTCGACAGCCGGA				866
Db	333	C-----CGCCCG				276
QY	867	TCAGGCGCGCGCGCTTCGATTCGAGCCATAGCTGTGCTGACGAGCGCGGATGCTGGGGA				926
Db	275	CGCCCGG				216
QY	927	ACGCGACCTTCGTATTCGTCCTCATTTCCGCGCGACGATGCGGCGACAGCGCGCGCACCATTT				986
Db	215	ACGG				156
QY	987	CGGCGGAAAGATCGGTGTCTGTGGCAGGACACAGGTGTGCTGTTCGAGGGGCGGAGCCGCG				1044
Db	155	GCGGGGGCGCGCGCGCGCGCGCGGGGGGGGGGGGGGGGGCGCGCGCGCGCGCGCGCGCGCG				96
QY	1047	CGTCAGAGCATCACGATTCGCGCGCATCCGGTCTGCGCGTTCGCGAACGCGCAGCGCGGATTCAGG				1104
Db	95	CGCGAGAAAGGGGGGGGGCGCGCGCGCGCGCGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG				36

[illegible]

[illegible][illegible]

QY 128 TCGATCAGCGCTGCGGAGCGGAGTAAACCGCTGACGAGCGGATAGCGAGCGTGGCG 187
DB 1013 GCGCGGCGGCG 954
QY 188 GGGGACCGCGGAGACGATCCGTTTCAGCAGCGCGGAGGAGCGGTGCGGATCCGCGGCA 247
DB 953 CCGGCG 894
QY 248 TCGATGCG 304
DB 893 GCG 835
QY 305 GCGATGCGATCCGCGGAGCTGCGCGGAGTAAAGCGAGCGGATATCCGTTGACGCGGTTGAAAC 364
DB 834 GCG 775
QY 365 AGCGCTGCG 417
DB 774 CCG 715
QY 418 GCGTATGCGGTCATGCG 477
DB 714 CCG 655
QY 478 CCGGATCGACCG 537
DB 654 CCG 595
QY 538 CAGATCG 563
DB 594 CCG 569

RESULT 6
AG171124/c 1798 bp DNA linear GSS 09-JAN-2002
LOCUS Pan troglodytes DNA, clone: RPA3-040F09.TJ, genomic survey
DEFINITION sequence.
ACCESSION AG171124
VERSION AG171124.1 GI:1670802
KEYWORDS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

REFERENCE
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
TITLE Totsuki, Y., Watanabe, H. and Sakaki, Y.
JOURNAL BAC end sequences of Library RPCI-43
REFERENCE Unpublished
AUTHORS
TITLE 2 (bases 1 to 1798)
JOURNAL Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totsuki, Y., Watanabe, H. and Sakaki, Y.
DIRECT SUBMISSION
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan
Tel: 81-45-503-9111, Fax: 81-45-503-9170
E-mail: chimpeseg@riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.

COMMENT
PRIMERS
Sequencing: TJ

LIBRARY
Vector : pBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
Location/Qualifiers
1.1798
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"

ORIGIN
Query Match 6.2%; Score 71; DB 9; Length 1798;
Best Local Similarity 41.0%; Pred. No. 6,9e-05;
Matches 463; Conservative 0; Mismatches 644; Indels 21; Gaps 5;
/clone="RPA3-040F09.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

QY 2 CATGCTCTCTCTGACAGAGGCGCGTTGCGGACAGGAGCGGAGCGGCTGCGGACAGCGGA 61
DB 1718 CUNGGAGGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1659
QY 62 ATGGGCGGCG 121
DB 1658 GNGCG 1599
QY 122 AAGCGCTGATCAGCGCGCTGCGGCGGAGCGGTGAAACCGCTGCGGAGCGGATAGCGAGCG 181
DB 1598 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1545
QY 182 TCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 241
DB 1544 GNGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1485
QY 242 GCGGATCGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
DB 1484 GCG 1425
QY 302 GCGCGGATGCGATCCGCGGAGCTGCGCGGCGGAGCGGAGTATCCGTTGAGCGGAGTGG 361
DB 1424 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1365
QY 362 AACGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 421
DB 1364 CCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1305
QY 422 ATGGCGTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481
DB 1304 GNGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1245
QY 482 GTCGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
DB 1244 CCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1188
QY 542 TCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601
DB 1187 CCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1128
QY 602 AGATGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661
DB 1127 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1068
QY 662 TCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 721
DB 1067 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1008
QY 722 GTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 781
DB 1007 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 948
QY 782 GAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 841
DB 947 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 892
QY 842 AGATGAGACCGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 901
DB 891 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 839
QY 902 GTCAGCGCGCGGAGATGCGGCGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 961
DB 838 CCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 779

[illegible]

Oy	195	GCGGAAACAGCATTCGGGTTCACAGACGCCGAAGAAGCGATTCGGATTCGGGCATGATG	254
Db	1125	GCGGGGGCCG - GCGGCCCGGGGGCCCCCGCGCGCGCCCGCGGCGCCCGGCCCGC	1067
Oy	255	CCACACC GGCA CCGCGCACCGGGGAGACGCGGTCTCAGATTGGCGCGCGCGAATGACAT	314
Db	1066	CTCCGCGCGCCCGCGGGGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1007
Oy	315	CCGCAACTGCGCGGCATTAGGCGACGGAATATCCGATGACGGGGTGGAAACAGCCCTGCC	374
Db	1006	CGGCCCCCCCCCGGGGGCGCGCGCGCGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCG	947
.Oy	375	CCAGCCCAAACGGGACAACGGCCCCCTGGGGTGATTCGCGCAGAAAGCTATGAGGTATGGG	434
Db	946	CGGCGCGGCGCCCCCGCGCGCGCGCGCGCGCGCGCGCG - CGCGCGCGCGGGGCCCG	888
Oy	435	CCAGCGCATGGGACAGATGCGCCTTTTCGCGCGCGCATCTCTGCGCGGTCAGACCCCGCC	494
Db	887	CGGTGCGGGCCGGGGGCGCGGGGCCCCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	828
Oy	495	TGGCGGCATAGTCAGCGACGCGCTTGCGCGCAGCGCGCATTCGTCAGATGCGCGCGTGC	554
Db	827	GCGGCGCGGGGCGCGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	768
Oy	555	TGTACGCGGTATCTTCGATCAGAAATCGGGTGGACTGAAGGCGACAGATGATGAAC	614
Db	767	COCGCCCCCCCCCGCCCTTCGACGCCAACCGGCGCGGGCGCGCGCGCGCGCGCGCGCG	708
Oy	615	GGTACCGCGCATCTGCGGGAACGHTCGGTCCATGATCATCGGGGCGTCGACGCATGGG	674
Db	707	GAGGCGCCCGGCGCGCGCGCGCGGGGGGGGGCGGGG --- GGGGCGCGCGCGCGCGGGG	652
Oy	675	GAGCGTCGGGTCTCGATCTCGACGCGCCACGAATTTCTGGAACCCACGATCAGATCGGGG	734
Db	651	GGGGGGGGGGGGGTTTTGTTCCCCCCCCCCCCCCCCCCCCCCCCCGGGGGGGGGGGGG	592
Oy	735	TCTCGACGACACACAGGGCGTGCATCACGACGACAGCCTCGATTCGCGAGCGGCTCGTCA	794
Db	591	GAGCGGCGGGCGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGGGGGGGCGCGCGCGCG	532
Oy	795	GCGTCGCGCCGCGTATCTTCGACGCGTTCGCGACATGCGTATTCACCGCACGATGACACCT	854
Db	531	GCGGGGCG	472
Oy	855	GCAGCAGCCGCATCAGCGCGCGCG 878	
Db	471	GCGGCGCGCGCGCGCGCGCGCG 448	

RESULT 9	AG030608/c			
LOCUS	AG030608	1189 bp	DNA	linear
DEFINITION	Pan troglodytes DNA, clone: PTB-003A18.R, genomic survey sequence.			
ACCESSION	AG030608			
VERSION	AG030608.1	GI:16557481		
KEYWORDS	GSS.			
SOURCE	Pan troglodytes (chimpanzee)			
ORGANISM	Pan troglodytes			
REFERENCE	Buxaryota, Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.			
AUTHORS	1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.			
TITLE	BAC end sequences of library PTB			
JOURNAL	Unpublished			
REFERENCES	2 (bases 1 to 1189)			
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpbees@cc.riken.go.jp, URL: http://hnp.gsc.riken.go.jp/),			

COMMENT

Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

Sequencing: M13Rev
LIBRARY

1. .1189

OPTIC

Query Match	6.1%;	Score 69.8;	DB 9;	Length 1189;
Best Local Similarity	47.8%;	Pred. No. 0.00013;		
Matches 261;	Conservative	0;	Mismatches 278;	Indels 7;
				Gaps 3;

Oy		5	GCTCTTCCTCAGCAAGGGGACGTTCCGAGCAGGAGCGCACGGCCTTCCGAACAAGGGAAATG	64
Db		826	GCGCGCGGGGGGCCGGGGGGGGCGCGGGGGCCCCGCCTGCTGCGGGGGCGGTGGGGCCGGGGGGGGGG	767
Oy		65	GGCGGGGCGTCCGGTGACGATTCGGAAGCGCGGTCCGGCCATGTCAAGCGCCCGCGCATAGAAG	124
Db		766	CGCGGGGCGGCGCGCCGGGGG---GGCGGGGCGNCCC CGGCGCCCGCGGGCGGGGCGCGCG	710
Oy		125	CGCTTCATTACGCGGCTGCGGCGAGGCGGTAGAACCTGCTGCACAGCGAATAGCGATCG	184
Db		709	CCCCCGGCGCGGGC--CCGNCGGCCGCGCCGCGCGCGCGCGCGCGCCCGCGCGCGCC	652
Oy		185	GGCGGGGAGCGCGCGGMAACAGCATTCGGTTCAAGCAGCGCGAGAAAGCGGTCCGATCCGGC	244
Db		651	GCGCGCTCGGCGCGCGCGGNCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC	592
Oy		245	CGATCGATTGCGCCAGCGCGCGCACCGCGCGCACGGGCGGACCGCGTGTCAAGTCCGCGGCC	304
Db		591	GCGGCGG	532
Oy		305	GCGATGCGCATTCGCGACCTGCGCGGCAATAGGGGACGCAATATCCGGTGACGGGGTGAAC	364
Db		531	CCGC	472
Oy		365	AGCGCTGCGCGCGCAAGCCGAACCGGACACCGGCCCTTGCGCGTGTGCGGCGCAAGAACCTATG	424
Db		471	GCGCGCGCGCGCGGGGGGCGCGNCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	412
Oy		425	GCGTCATGAGGCGACGCGGATGAGGCGAGGATCCCTTTTCGCGCGCATCTTCGTCGCGGTC	484
Db		411	GNGCGCGCGCGCGGGCGGAGCGNCCCGCGCGCNCCGCGGGGGGCGCGGGGGCGGGGCCNNC	352
Oy		485	CAGGCCCGCTGCGCGCATAGT--CGAGCGACGCTGCGCGCAGCGCGGCATGCTCCAGAT	542
Db		351	NCCCCCGNCCCCCGNCAAACAAGCAGCAAACTATCTCTGNCGCGCGCGCCCCCCCC	292
Oy		543	CGCGCG	548
Db		291	CGCGCG	286

RESULT 10	CG756569	1406 bp	DNA	linear	GSS 24-OCT-2003
LOCUS	CG756569				
DEFINITION	CG756569	P051.4-B06.zb Ppa EcORI BAC Library	Pristionchus pacificus	genomic,	
ACCESSION	CG756569	genomic survey sequence.			
VERSION	CG756569.1	GI:37984265			

KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1406)
AUTHORS Srinivasan,J., Sinz,M., Jesse,T., Wiggers-Pereboile,L., Jansen,K.,
Bunjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22815951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel.: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES
source location/Qualifiers
1..1406
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 6.1%; Score 69.6; DB 9; Length 1406;
Best Local Similarity 44.6%; Pred. No. 0.00014;
Matches 424; Conservative 0; Mismatches 518; Indels 9; Gaps 4;
179 CGGTGGGCGGCGAGCGCGGAAAGCATCCGGTTTCAGAGCGCGAGAAAGCGTTCGCGA 238
294 CCGGCG 353
239 TC CGCGCGCATGATGGCG 298
354 CCG 413
239 CCGCGCGCGCATGCGCATCG 358
414 CCG 473
359 TGGAAAGCG 418
474 GAGGCG 533
419 CCTATGGCGTCAATGGCG 478
534 GCG 591
479 CCGGTTCAGCG 538
592 CCG 651
539 AGATCG 598
652 GCG 711
599 AGCAGATGATGAAGCGGTACCGGTCACTCGGGAACGCGCGGTCACTGATCATCGAG 658
712 GCG 766
659 CGCTGACCGCATGGGCGCGCGGTGCTGATCTCGAGCGCCACGAATTTCTGGAAGCC 718
767 CCG 826

QY 719 ACGGTCAAGTG-CGGGTCCTGACAGGACCAACGCGGCGCTTCATCAAGCAGGAGCTTCAT 777
DB 827 CCG 886
QY 778 CCGGAGCGCGTCCGTCAGCGGTTCGCGCGGTATCGTCAGCGCGCGCATCGATTCCA 837
DB 887 GCG 946
QY 838 CCGGAGATGCAACCGTTCAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
DB 947 GCG 1006
QY 898 TGTGCTCAAGCG 957
DB 1007 CCG 1065
QY 958 ACGATGGCG 1017
DB 1066 GCG 1125
QY 1018 GGTGTGCTGTCG 1077
DB 1126 CCG 1185
QY 1078 GCGGTGCGGAACGCGCAACCGCGATGCGCGCACCGGACGCGCGCGCGCGCGCGCGCG 1128
DB 1186 CCG 1236

RESULT 11
LOCUS AG036223 938 bp DNA linear GSS 01-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-012B18.F, genomic survey sequence.
ACCESSION AG036223
VERSION AG036223.1 GI:16563096
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 938)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
E-mail:chimbases@sc.riken.go.jp, URL:http://npg.gsc.riken.go.jp/
Tel:81-45-503-9111, Fax:81-45-503-9170
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
source location/Qualifiers
1..938
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-012B18.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC library"

ORIGIN

Bullis, D., Snell, J., Miguel, T., Hutchinson, D., Kimerly, B.,
Mitzel, T., Katagiri, F., Glazebrook, J., Law, M., and Goff, S. A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)

JOURNAL
MEDLINE
PUBMED

12468722

Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com

ABRC Stock Number CS830561; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: T-DNA tagged.

FEATURES
source

1..929
location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_694_B04.v1"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a
modified TAIL-PCR strategy"

ORIGIN

Query Match 5.9%; Score 68.2; DB 9; Length 929;
Best Local Similarity 47.9%; Pred. No. 0.0028;
Matches 210; Conservative 0; Mismatches 227; Indels 1; Gaps 1;

QY 57 GCGGATGGGCGGCGGCTCCGCTGACGATCGAAGCCGTCGCGCCATGTCAGCGCCCG 116
DB 407 GGGGGGGGGGGGGGGGGGAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 466
QY 117 CATGAAGGCTGATCAGCGGCTGCGGCGAGCGGTGAACCGCTGCAAGCAGCGCATAGC 176
DB 467 GGGGGGGGGGGGGGNNNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 526
QY 177 GACGCTCGGCGGCGGCGGCGGAGACAGATCCGTTTCAGCAGCCGAGGAACCGTCCG 236
DB 527 GGGGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 586
QY 237 GATCGCGCGATGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 296
DB 587 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 646
QY 297 CGCGCGCGCGATGCGGATCGGACCTGCGCGGCGATAGGGGAGCGAGCATATCCGATGACG 356
DB 647 GGGGGGGGGGGGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 706
QY 357 GGTGAACAAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 416
DB 707 GCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 765
QY 417 AGCCTATGGCGTCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 476
DB 766 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 825
QY 477 GCCCGATCCAGCCCGCC 494
DB 826 GCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 843

RESULT 14
AG159205/C 896 bp DNA linear GSS 09-JAN-2002
LOCUS AG159205
DEFINITION Pan troglodytes DNA, clone: RP43-024K15.T7, genomic survey
SEQUENCE
ACCESSION AG159205
VERSION AG159205.1 GI:1668883

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL

GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H., and Sakaki, Y.
BAC end sequences of library RPCT-43
Unpublished
2 (bases 1 to 896)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H., and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9111, Fax: 81-45-503-9170
E-mail: chimpanzee@sc.riken.go.jp, url: http://ngp.gsc.riken.go.jp/
Clones are derived from the chimpanzee BAC library RPCT-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.

COMMENT

PRIMERS
Sequencing: T7

FEATURES
source

LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
location/Qualifiers
1..896
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-024K15.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCT-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 5.9%; Score 68; DB 9; Length 896;
Best Local Similarity 47.3%; Pred. No. 0.003;
Matches 259; Conservative 0; Mismatches 286; Indels 3; Gaps 2;

QY 5 GCTCTCTCTTCGACGAGGAGGCGTTTCGCGGCGAGGCGGCGGCGGCGGCGGCGGCGG 64
DB 844 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 785
QY 65 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 124
DB 784 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 725
QY 125 GCGTATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
DB 724 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 665
QY 724 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 665
DB 183 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
QY 664 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 605
DB 243 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
QY 604 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 545
DB 302 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 361
QY 544 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 485
DB 362 AACAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421
QY 484 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 425
DB 422 ATGCGCTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 22:57:55 ; Search time 826.57 Seconds
(without alignments)
8628.994 Million cell updates/sec

Title: US-10-695-980-1_COPY_5794_6942

Perfect score: 1149

Sequence: 1 tcatgctctctctcgcagca.....atcagcagatcatgctcat 1149

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubna/US05_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubna/US04_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubna/US03_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubna/US02_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubna/US01_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubna/US00_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubna/US05_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubna/US04_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubna/US03_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubna/US02_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubna/US01_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubna/US00_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
23: /cgn2_6/ptodata/1/pubna/US05_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubna/US04_PUBCOMB.seq:*
25: /cgn2_6/ptodata/1/pubna/US03_PUBCOMB.seq:*
26: /cgn2_6/ptodata/1/pubna/US02_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1149	100.0	1149	9	US-09-547-267-8
2	1149	100.0	8625	10	US-09-920-923-1
3	1149	100.0	8625	18	US-10-695-980-1
4	1149	100.0	11233	10	US-09-920-923-27
5	1149	100.0	11233	18	US-10-695-980-27
6	248.8	21.7	1164	21	US-10-810-733-7
7	248.8	21.7	9127	21	US-10-810-733-20

C 8	236.4	20.6	1185	14	US-10-166-037-2
C 9	228.6	19.9	1170	20	US-10-808-807-5
C 10	228.6	19.9	5632	22	US-10-997-844-6
C 11	228.6	19.9	8814	20	US-10-808-807-18
C 12	228.6	19.9	8814	22	US-10-997-844-11
C 13	213.4	18.6	12753	18	US-10-041-018-19
C 14	213.4	18.5	1167	20	US-10-804-677-5
C 15	213	18.5	8675	20	US-10-804-677-18
C 16	209.8	18.3	1149	10	US-09-941-947A-29
C 17	209.8	18.3	1149	15	US-10-218-118-5
C 18	209.8	18.3	1149	18	US-10-363-567-29
C 19	209.8	18.3	1149	18	US-10-466-656-3
C 20	209.8	18.3	1149	19	US-10-735-019-5
C 21	209.8	18.3	1149	20	US-10-734-778-5
C 22	209.8	18.3	1149	20	US-10-735-442-5
C 23	209.8	18.3	1149	21	US-10-848-307-3
C 24	209.8	18.3	1149	21	US-10-886-906-5
C 25	209.8	18.3	3611	22	US-10-997-844-36
C 26	209.8	18.3	8609	19	US-10-735-019-28
C 27	209.8	18.3	8609	20	US-10-734-778-40
C 28	209.8	18.3	8609	20	US-10-734-936-37
C 29	209.8	18.3	8609	20	US-10-735-442-64
C 30	209.8	18.3	8609	21	US-10-886-906-54
C 31	187.4	16.3	1167	20	US-10-808-979-5
C 32	187.4	16.3	6999	20	US-10-808-979-18
C 33	174.8	15.2	7494	19	US-10-474-536-45
C 34	174.8	15.2	8547	19	US-10-474-536-46
C 35	65	5.7	125401	19	US-10-203-295-35
C 36	64.6	5.6	2209	18	US-10-425-114-23000
C 37	64.4	5.6	1041	19	US-10-425-963-8597
C 38	63.2	5.5	2265	18	US-10-425-114-26596
C 39	63.2	5.5	2451	18	US-10-425-114-24537
C 40	63	5.5	2688	20	US-10-425-115-75178
C 41	62	5.4	1859	17	US-10-282-122A-14725
C 42	60	5.2	2992	17	US-10-369-423-31508
C 43	59.6	5.2	88421	9	US-09-976-055-1
C 44	59.2	5.2	1684	19	US-10-437-963-81995
C 45	59.2	5.2	1998	9	US-09-813-467-1

ALIGNMENTS

RESULT 1
US-09-547-267-8/C
; Sequence 8, Application US/09547267
; Patent No. US20020147371A1
; GENERAL INFORMATION:
; APPLICANT: Hohmann, Hans-Peter
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tessier, Michel
; APPLICANT: van Loon, Adolfinus
; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/547,267
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/660,645
; FILING DATE:

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Pokias, Bruce A.
/ REGISTRATION NUMBER: 32,748
/ REFERENCE/DOCKET NUMBER: RAN 6002/170
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (201) 235-5801
/ TELEFAX: (201) 235-2363
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1149 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-09-547-267-8

Query Match      100.0%; Score 1149; DB 9; Length 1149;
Best Local Similarity 100.0%; Pred. No. 9.7e-282;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGCTCTCTCTGACAGAGGGGGGCTTCGAGGACAGGCGGACCGCTTGACAGCGG 60
DB 1149 TCATGCTCTCTCTGACAGAGGGGGGCTTCGAGGACAGGCGGACCGCTTGACAGCGG 1090
QY 61 AATGGGCGGGCGTCCGGTGACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCGGCATTA 120
DB 1089 AATGGGCGGGCGTCCGGTGACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCGGCATTA 1030
QY 121 GAAGCGCTGATCAGCGGCTGCGGCGAGGGGTAGAAACCGCTCAGAGCGCATAGCAGG 180
DB 1029 GAAGCGCTGATCAGCGGCTGCGGCGAGGGGTAGAAACCGCTCAGAGCGCATAGCAGG 970
QY 181 GTCGGGCGGGACCGCGGAAACGATCCGGTTTCAGCAACCCGACAGAAACGGTCCGATC 240
DB 969 GTCGGGCGGGACCGCGGAAACGATCCGGTTTCAGCAACCCGACAGAAACGGTCCGATC 910
QY 241 CGCGCGATGATGAGCCGACCGCGCACCGCGCGACCGGCGGACCGCGTGTCAAGTCCG 300
DB 909 CGCGCGATGATGAGCCGACCGCGCACCGCGCGACCGGCGGACCGCGTGTCAAGTCCG 850
QY 301 CGCGCGATGATGAGCCGACCGCGCACCGCGCGACCGGCGGACCGCGTGTCAAGTCCG 360
DB 849 CGCGCGATGATGAGCCGACCGCGCACCGCGCGACCGGCGGACCGCGTGTCAAGTCCG 790
QY 361 GAACAGCGCTGCGGCGGACCGCGCACCGCGCGGCGGCGTGTCAAGTCCGCGCAAGCC 420
DB 789 GAACAGCGCTGCGGCGGACCGCGCACCGCGCGGCGGCGTGTCAAGTCCGCGCAAGCC 730
QY 421 TATGGCGTATGAGGCGGACGATGAGGCGGATGCCCTTTTCGCGCGCATCTCTGCC 480
DB 729 TATGGCGTATGAGGCGGACGATGAGGCGGATGCCCTTTTCGCGCGCATCTCTGCC 670
QY 481 GGTCCAGCGCGCTTGGCGGACATGTCAGCGACCGCTCGCGGCGGCGGCGGCGGCGG 540
DB 669 GGTCCAGCGCGCTTGGCGGACATGTCAGCGACCGCTCGCGGCGGCGGCGGCGGCGG 610
QY 541 ATCGCGCGCGCTGCTGAGCGGATCTCGATCAGAGTGGGGTGGAGTGAAGGCGAG 600
DB 609 ATCGCGCGCGCTGCTGAGCGGATCTCGATCAGAGTGGGGTGGAGTGAAGGCGAG 550
QY 601 CAGATGATGAAGCGGTTACCGCTCATCTGCGGAACGGTCCGCTCATGATCATCGGG 660
DB 549 CAGATGATGAAGCGGTTACCGCTCATCTGCGGAACGGTCCGCTCATGATCATCGGG 490
QY 661 CTCGAGCGCATGGGGGGCGTCCGTCTCGATCTGAGAGCCCAAGAAATTTCTGAAACCCAC 720
DB 489 CTCGAGCGCATGGGGGGCGTCCGTCTCGATCTGAGAGCCCAAGAAATTTCTGAAACCCAC 430
QY 721 GGTCAAGGTGAGGGGTCTCGACGCGCACGCGGGCGTATCATCAGCAGACCGCTCATCG 780
DB 429 GGTCAAGGTGAGGGGTCTCGACGCGCACGCGGGCGTATCATCAGCAGACCGCTCATCG 370
QY 781 CGAGCGGTCTGACGCGTCCGCGCGGATGTCAGCGTCCGCGCATGCGTATTCCACCG 840
```

```
DB 369 CGAGCGGTCCGTCAAGCGGTGCGCGCGGTATCGTCAAGCGTCCGACATGCGATTCACCG 310
QY 841 CAGATGACACCGCTTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 309 CAGATGACACCGCTTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 250
QY 901 CGTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 249 CGTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 190
QY 961 AATGGCGCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 189 AATGGCGCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 130
QY 1021 GTGCTGTCCGAGGGGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 129 GTGCTGTCCGAGGGGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 70
QY 1081 GTCGGAACGGCAAGCGGATCAGGCGCACCGGACCGGCGGCGGCGGCGGCGGCGGATC 1140
DB 69 GTCGGAACGGCAAGCGGATCAGGCGCACCGGACCGGCGGCGGCGGCGGCGGCGGATC 10
QY 1141 ATGCTCAT 1149
DB 9 ATGCTCAT 1

RESULT 2
US-09-920-923-1
/ Sequence 1, Application US/09920923
/ Publication No. US2003002273A1
/ GENERAL INFORMATION:
/ APPLICANT: Pasamontes, Luis
/ APPLICANT: Tsygankov, Yuri
/ TITLE OF INVENTION: Improved Fermentative Carotenoid Production
/ FILE REFERENCE: Improved Fermentative Carotenoid
/ CURRENT FILING DATE: 2001-08-02
/ PRIOR APPLICATION NUMBER: 08/980,832
/ PRIOR FILING DATE: 1997-12-01
/ NUMBER OF SEQ ID NOS: 66
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 8625
/ TYPE: DNA
/ ORGANISM: Flavobacterium sp. R1534
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (8348)..(8349)
/ NAME/KEY: unsure
/ LOCATION: (8539)..(8540)
/ NAME/KEY: unsure
/ LOCATION: (8581)
/ NAME/KEY: unsure
/ LOCATION: (8590)
/ NAME/KEY: unsure
/ LOCATION: (8592)
/ NAME/KEY: unsure
/ LOCATION: (8602)..(8604)
US-09-920-923-1

Query Match      100.0%; Score 1149; DB 10; Length 8625;
Best Local Similarity 100.0%; Pred. No. 8.3e-282;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGCTCTCTCTGACAGAGGGGGGCTTCGAGGACAGGCGGACCGCTTGACAGCGG 60
DB 5794 TCATGCTCTCTCTGACAGAGGGGGGCTTCGAGGACAGGCGGACCGCTTGACAGCGG 5853
QY 61 AATGGGCGGGCGTCCGGTGACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCGGCATTA 120
DB 5854 AATGGGCGGGCGTCCGGTGACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCGGCATTA 5913
```


Db	6094	CGCCGCATGGCATTCGGCACTTGCGCGGCATAGGGACAGCAATATCCGGTGAAGGGGTG	6153
Qy	361	GAACAGCCTTACCCTCCAGCCCAACCGGCAACCGCCCTTGCGCGTGTGCGCCAGAAACC	420
Db	6154	GAACAGCCTTACCCTCCAGCCCAACCGGCAACCGCCCTTGCGCGTGTGCGCCAGAAACC	6213
Qy	421	TATGGCGTCATAGGGCCAGCGCGATGGGAGATGCCCTTTTCGGCGCGCATCTCTGCC	480
Db	6214	TATGGCGTCATAGGGCCAGCGCGATGGGAGATGCCCTTTTCGGCGCGCATCTCTGCC	6273
Qy	481	GGTCAGCGCCCGCTGGGCGCATATGTCAGGAGACGCTTGCGCGCCAGCGCGCATGCTCCAG	540
Db	6274	GGTCAGCGCCCGCTGGGCGCATATGTCAGGAGACGCTTGCGCGCCAGCGCGCATGCTCCAG	6333
Qy	541	ATCCGCGCGCTGCTGTATAGCGCGTATCTTCGATCAGATGCGGGTGGGACTGAAGGGCAG	600
Db	6334	ATCCGCGCGCTGCTGTATAGCGCGTATCTTCGATCAGATGCGGGTGGGACTGAAGGGCAG	6393
Qy	601	CAGATAGATGAAGCGGTACCCGTCATCTGCGGAAACGGTCGCGTCATGATCATCGGCG	660
Db	6394	CAGATAGATGAAGCGGTACCCGTCATCTGCGGAAACGGTCGCGTCATGATCATCGGCG	6453
Qy	661	CTCCACGCGCATAGGGGGGCGTGGTTCGATCTCCACGCGCCAGTCCGAATTTCTGGAAACCCAC	720
Db	6454	CTCCACGCGCATAGGGGGGCGTGGTTCGATCTCCACGCGCCAGTCCGAATTTCTGGAAACCCAC	6513
Qy	721	GGTCAGGTGCGGGGTCTCGACGCGCACACGCGGCGTCGATTCACGACGAGCGCTGATCCG	780
Db	6514	GGTCAGGTGCGGGGTCTCGACGCGCACACGCGGCGTCGATTCACGACGAGCGCTGATCCG	6573
Qy	781	CGAGCGGTCCGTCAAGCGTGCAGCGCGCGGTATCGTCAAGCGTGCAGCATGCTATTTCCACCG	840
Db	6574	CGAGCGGTCCGTCAAGCGTGCAGCGCGCGGTATCGTCAAGCGTGCAGCATGCTATTTCCACCG	6633
Qy	841	CAGATTCGACACCTTCGACACAGCCGATAGGCGCGCGCCCTCGATTCAGAGCCATAGCCTGT	900
Db	6634	CAGATTCGACACCTTCGACACAGCCGATAGGCGCGCGCCCTCGATTCAGAGCCATAGCCTGT	6693
Qy	901	CGTCAGCGCGCGCGCATAGTGGGGGAAACGCACTCTGATCCGTCATTCGCGCGCGACG	960
Db	6694	CGTCAGCGCGCGCGCATAGTGGGGGAAACGCACTCTGATCCGTCATTCGCGCGCGACG	6753
Qy	961	AATGGGCGACAGGCGCGCGCAGCCATTTCGGGCGAAAGATCCGTGTCTGGCGACGACGT	1020
Db	6754	AATGGGCGACAGGCGCGCGCAGCCATTTCGGGCGAAAGATCCGTGTCTGGCGACGACGT	6813
Qy	1021	GTCGCTGTCTCCAGAGGGGCCGAGACCGCGCTCGAGCATTCACATTCGCGCATCCGGTCTGCG	1080
Db	6814	GTCGCTGTCTCCAGAGGGGCCGAGACCGCGCTCGAGCATTCACATTCGCGCATCCGGTCTGCG	6873
Qy	1081	GTCCGGAACGGGCAACGCGGATCAGCGGACCGGACGCGCGCGCGCGCGCATCAGCAATC	1140
Db	6874	GTCCGGAACGGGCAACGCGGATCAGCGGACCGGACGCGCGCGCGCGCGCATCAGCAATC	6933
Qy	1141	ATGGCTCAT 1149	
Db	6934	ATGGCTCAT 6942	
RESULT 4			
US-09-920-923-27			
; Sequence 27, Application US/09920923			
; Publication No. US2003022273A1			
GENERAL INFORMATION:			
; APPLICANT: Pasamontes, Luis			
; APPLICANT: Tsygankov, Yuri			
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production			
; FILE REFERENCE: Improved Fermentative Carotenoid			
; CURRENT FILING DATE: 2001-08-02			
; CURRENT APPLICATION NUMBER: US/09/920,923			
; PRIOR APPLICATION NUMBER: 08/980,832			
; PRIOR FILING DATE: 1997-12-01			

[illegible]

```

RESULT 5
US-10-695-980-27
; Sequence 27, Application US/10695980
; Publication No. US20040058410A1
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tszygankov, Yuri
; TITLE OF INVENTION: Fermentative Carotenoid Production
; FILE REFERENCE: 15464 US (C38435/125944)
; CURRENT APPLICATION NUMBER: US/10/695,980
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/920,923B
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentin Version 3.1
; SEQ ID NO 27
; LENGTH: 11233
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Plasmid pzea4
US-10-695-980-27

```

[illegible]

;; PRIOR FILING DATE: 2003-07-17
;; PRIOR APPLICATION NUMBER: US 60/527,083
;; PRIOR FILING DATE: 2003-12-03
;; NUMBER OF SEQ ID NOS: 22
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 7
;; LENGTH: 1164
;; TYPE: DNA
;; ORGANISM: Pantoea stewartii DC413
US-10-810-733-7

Query Match 21.7%; Score 248.8; DB 21; Length 1164;
Best Local Similarity 54.2%; Pred. No. 1.5e-53;
Matches 605; Conservative 0; Mismatches 487; Indels 24; Gaps 4;

QY 42 GCACGGCCCTGCACAGCGGAATGGCGGCGTCCGGTGAAGATGGCAAGCCGGTGGCCA 101
DB 1120 GCACGGCCCGCCAGCACCGGAGACCGGCGCTTGCGCTTAAGAAATGCGCGCGCGTGGCCA 1061
QY 102 ATGTACAGCGCCCGCGCATGAAGCGCTCGATCAGCGGCTGCGGCAAGCGGTAGAACCGCT 161
DB 1060 GGGTCAGCGCGTCCGGGTAAACCGGGCGATCAGCCCTCGGGCAGCGCGTAAAGCGCT 1001
QY 162 GCACGAGCGGATGCGAGCGGTGGGCGGCGAGCCGCGGAACAGCATCCGTTTACAGACGC 221
DB 1000 GCATTACGCGCAGCGCCGATCGCCCTCGGCGCGCCAGGAACAGCATGCGGTTAAGCATGC 941
QY 222 GCAGGAGAGGGTCCGGATCCGGCGCATGATGGGCCAGCGCGGCAACCGGCGAGCGGG 281
DB 940 GGAAGAAAGCCCTGTTTAAACGACGCTGCTCGGCAAGGGGTGAATACCGCGCGCAGCG 881
QY 282 ACGCGGCTGTCAAGTGCAGC---GCCGCGATGGCATCCGCGACTTCCGCGGATAGGGA 338
DB 880 CCTCAGGGGAGAAATCGCGCTCTCGGGCGAGGGCGTCCGCGACGCTGCGCGCAGCGGA 821
QY 339 GCGAATATCCGTGACGGGGTGAACAGCCCTGCCCCAGCCCAACCGGCAACCGCCCT 398
DB 820 GGGAGTAGCCGGTCTGGGGTGAAGAAAGCCCGCGCGCAGCGCGCAGG-----CAGG 770
QY 399 GGGCGGTGTGGCGCCAGAAAGCCATGGCGTATGGGCCAGCGCGCATGGGCGAGATGCCCC 458
DB 769 GGGCGCGATGATGCCAGAAAGCGGTCAAAATGCCCGCCAGCGTAATCGGCGAGCGCCCG 710
QY 459 T---TTCCGCGCGCATCTCTGCGCCGATCCAGCGCCCGCTGCGCGCATAGTCCAGCGAGC 515
DB 709 GCTCTTCGGGAGAGACGCTCTCAGGTGCGACGCGCTGGGGGTGGGTAGTGGCAGATAT 650
QY 516 CCTGCGCCAGCGCGCATGCTTCAGTGCCTGCGCTGCTGTAAGCGGATCTCTGATTA 575
DB 649 TCTGCCGCGCTGCGCGCGCTGCGCGCAGGCGCTGTTAATATATGTTGATCTTCATCA 590
QY 576 GGAATGGGGTGGGATGAAAGGCGAGCATGATGAAAGGGTAAACCGGTCAATCTGGGAA 635
DB 589 GGAAGAGCGGTCCGCGAGGCGAGGGATTAAGCAGAGGATAGCCCCCTGTGATATCA 530
QY 636 CGGTCCGCTCATGATCATCGGCGCTCGACGCGCATGGGAGGCGTCTCGATCTCGA 695
DB 529 CGGCGCGCTCATCAAAATGCGCCCTTCAGGTGAGCGGCTGAGGCGCACTCCT 470
QY 696 CGCGCAGAAATTTCTGAAACCCACGCTGAGGTGCGGGGTCTTCGACGCGACCAACGCGGCT 755
DB 469 GACCGCAAAAGATGAAAGCCATCTGAGAGGGCGGCGTGGGGCTGATAGCGCGCGCGCT 410
QY 756 CGATACGCGAGGAGCGCTGATCCGAGACCGTCCCTCAAGCGTCCGCGCGGATATCTCA 815
DB 409 CAATAGCGCGCGCGCGCGAGCATTTGCGCTGGCGAGCTGCAAGTATCGGGCGCGC 350
QY 816 GGTGCGCAGCATGCGATTTCACCGGAGATCGACCT-----GACGAGCGCGCA 866
DB 349 TGTGCGCAGCGGTCTGTGTTAAGCGCGCGGTGCGCGCAGCGTCCGCGCGCAAGCACTCT 290
QY 867 TGAAGCGCGCGCTGATGAGCCATGACCTGTCTCTGACGCGCGCGCAATGCTGGGAA 926

DB 289 CAAGCGCGCGGAGTGAACGCAAAATATGCGCTGTTAAGCTGCGCGCTCAGCGCGGAA 230
QY 927 ACGCACTCTCATCTGATCCGTCATTGCGCGCAGCATGGGAGACAGCGCGCCAGCCATT 986
DB 229 AGCGTACCTCTGATGCG 170
QY 987 CGGCGGAAAGATCCGTGTCTGTGCGGAGACAGAGTGTCTGTGTCGAGGGGCGCGACCGC 1046
DB 169 CGGCGCTGATCTCTTCTGTGAAGACACAGTGTGATTGCGCACCGCGCGTGGCGTGG 110
QY 1047 CGTCAAGATACAGATGCGCGCATCCGGTCTGCGGTGCGGAAACGCGCAGCGCGATCAGG 1106
DB 109 CATCAAGACCAAGACGCGCGAGCTGCGCTGAGGCGCTTTACCGCGCAGCGCAGACGC 50
QY 1107 CACCGACAGCCCGCGCGCGCGCGATCAGCATATC 1142
DB 49 CTTTCCAGACCGCGCGCGCGCGAGCAACAGCATATC 14

RESULT 7
US-10-810-733-20/c
; Sequence 20, Application US/10810733
; Publication No. US20050014219A1

;; GENERAL INFORMATION:
;; APPLICANT: E.I. du Pont de Nemours and Co., Inc.
;; APPLICANT: Cheng, Qilong
;; APPLICANT: Tao, Luan
;; APPLICANT: Sedkova, Natalia
;; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
;; FILE REFERENCE: CL2385 US NA
;; CURRENT APPLICATION NUMBER: US/10/810,733
;; CURRENT FILING DATE: 2004-03-26
;; PRIOR APPLICATION NUMBER: US 60/488,183
;; PRIOR FILING DATE: 2003-07-17
;; PRIOR APPLICATION NUMBER: US 60/527,083
;; PRIOR FILING DATE: 2003-12-03
;; NUMBER OF SEQ ID NOS: 22
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 20
;; LENGTH: 9127
;; TYPE: DNA
;; ORGANISM: Pantoea stewartii DC413
US-10-810-733-20

Query Match 21.7%; Score 248.8; DB 21; Length 9127;
Best Local Similarity 54.2%; Pred. No. 1.3e-53;
Matches 605; Conservative 0; Mismatches 487; Indels 24; Gaps 4;

QY 42 GCACGGCCCTGCACAGCGGAATGGCGGCGTCCGGTGAAGATGGCAAGCCGGTGGCCA 101
DB 6138 GCACGGCCCGCCAGCACCGGAGACCGGCGCTTGCGCTTAAGAAATGCGCGCGCGTGGCCA 6079
QY 102 ATGTACAGCGCCCGCGCATGAAGCGCTCGATCAGCGGCTGCGGCAAGCGGTAGAACCGCT 161
DB 6078 GCGTAGCGCGTCCGCGTAAGACCGGAGATAGCCCTCGGCGAGCGCGTAAGAGCGCT 6019
QY 162 GCAGCAGCGATGAGGAGCGGTGGGCGGCGAGCCGCGGAAGAGCATCCGTTTACAGACGC 221
DB 6018 GCATTACGCGCAGGCGCGGATGCGCTGCGCGCGCGAGGAACAGCATGCGGTTAAGCATGC 5959
QY 222 GCAGGAGCGGTCCGGATCCGGCGCATGATGGGCCAGCGCGCGCAACCGCGAGCGGGCGG 281
DB 5958 GGAAGAAAGCCCTGTTTACGCGCAGCGTGCCTGCGCAAGCGGTGAATGACGCGCGCGAGCG 5899
QY 282 ACGCGGTGTCAAGTTCGCGC---GCCGCGATGAGCATCCGCGACTTCCGCGGATAGGGA 338
DB 5898 CTTCAAGGAGAGAAATGCTCCCTCTCTGCGGAGGCGTCTGCGCAGCGTCCGCGCAGCGGA 5839
QY 339 GCGAATATCCGTGACGGGGTGAACAGAGCCCTGCCCCAGCCCAACCGGCAACCGCCCT 398
DB 5838 GGGAGTAGCCGCTGTGGGGTGAAGAAAGCCCGCGCGCGAGCGCGCAGG-----CAGG 5788
QY 399 GCGCGTGTGTGCGCGCAAGACCTATGCGCTATGGGCGCAGCGCGATGGGCGAGATGCC 458

Db 5787 GGAGCGGATGATGTCGAGAGCGGTTCGAAATCGCCCGCGAGCTAATCGGAGCGCCCG 5728
Qy 459 T---TTGCGCCCGCATCTCTGCGCGGTTCAGAGCCCGCTGCGGCGATATCCAGGAGC 515
Db 5727 GCTTCTGCGCGCAGAGGTCTTCAGTCCAGCCCTGCGCGGTGCGGTAGTTCAGATAT 5668
Qy 516 CTGCGCCAGCGCGCATCTGTCAGATCGCGCGGTCTGATGCGCGATCTCTGATCA 575
Db 5667 TCTGCGCGCGCTGCGCGGTTCGCGAGAGCGGTCTTAATATAGTGGATCTTCAATCA 5608
Qy 576 GGATGCGGTGGAGTCGAAAGGAGAGATAGATGAAAGCGGATCCGTCATCTGCGGA 635
Db 5607 GCAGAGCGCGTCGCGAGAGCGGAGAGGATAGAGAGGATAGCCCGCTGCTGATCA 5548
Qy 636 CGGTCCCGCTTCATGATCATCTGCGCGGTTCAGAGCGGTGCGGCGGTCTGATCTCA 695
Db 5547 CGGCGCGGTTCATCAATTCGCGCGGTCTGCGAGTGAAGCGGTCTGAGCGCGCTCT 5488
Qy 696 CGCCCAAGATTTCTGGAACCCAGAGGTGAGTGGCGGTCTGAGCGGAGCAAGCGGCGT 755
Db 5487 GATCGCAAAAGCTGAAAGCCATCTGAGGCGCGGTCTGAGGCTGATAGCCCGCGGT 5428
Qy 756 CGATCAAGAGAGAGCGGTCTGATCGCGAGCGGTCTGAGCGCGGTCTGATCTGCA 815
Db 5427 CATGACGCGCGCGCGCGGAGCACTTCCGCGTGGAGCTGAGTATCGGCGCG 5368
Qy 816 GCGTCGCGAGATCTGATTCATCCGAGATGACACCTT-----GAGAGCGCGCA 866
Db 5367 TGTGCGCGAGCGGTCTGATTAAGCGCGAGCGGTCTGAGCGCGGTCTGAGCGCGCTCT 5308
Qy 867 TCAGCGCGCGCGGTCTGATGAGCACTGATCTGTCGAGCGCGCGGATGCTGCGGA 926
Db 5307 CAAGCGCGCGGAGAGTACGCAAAATAGCGCTGTTAGCTGCGCGGTGACGCGCGGA 5248
Qy 927 ACGCGACTCTGATCTGATCTGATTCATTCGCGCGAGCAAGTGGGCGCGCGCGCATT 986
Db 5247 AGGTACTCTGATGAGTGGCGCGAGCGGTGCGCGAGCGCGCGCATCTGATGCTGG 5188
Qy 987 CGGCGCAAAAGATCTGATCTGATGAGAGCAAGTGGTCTGATGAGGCGCGCGCGCG 1046
Db 5187 CGGCGGTGATCTCTTCTGTAAGCACCGAGGTGATGTCACCGCGGTGCGCGTGG 5128
Qy 1047 CGTCAAGCATCAAGTGGCGCGATCGGCTGCTGCGGTGCGGAAGCGAGCGGATGAGG 1106
Db 5127 CATCAAGCAAGCAAGCGAGCTGCGCTGAGCGCTTCAAGCGAGCGAGCGAGCG 5068
Qy 1107 CACCGAGAGCGCGCGCGCGCGATGAGATCAT 1142
Db 5067 CGTTGCGCAGAGCGCGCGCGAGCAAGAGATCAT 5032

RESULT 8
US-10-166-037-2/c
; Sequence 2, Application US/10166037
; Publication No. US20030087337A1
; GENERAL INFORMATION:
; APPLICANT: GIRAUD, Eric
; APPLICANT: HANNIBAL, Laure
; TITLE OF INVENTION: Isolated Carotenoid biosynthesis gene cluster involved
; TITLE OF INVENTION: in canthaxanthin production and application thereof
; FILE REFERENCE: 1721-55
; CURRENT APPLICATION NUMBER: US/10/166, 037
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/297, 272
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: cxy
US-10-166-037-2

Query Match 20.6%; Score 236.4; DB 14; Length 1185;
Best Local Similarity 54.0%; Pred. No. 2, 1e-50;
Matches 589; Conservative 0; Mismatches 471; Indels 30; Gaps 4;
Qy 78 TCAGATGCGAAGCGGTGCGGCAATGTCAGGCGCGCGCGATAGAGCGCTGATCAGG 137
Db 1102 TGAAGAGCGGAGCTTTGCTCTGCGGTGATCCGGCGCGCTGAGAGCGCTGATCAGG 1043
Qy 138 GCTGCGGAGGCGGTAGAACTGCTGAGAGCGATAGGAGCGTTCGCGCGCGAGCGCG 197
Db 1042 CCGTATGAGAGCGTAAATCGTGCAGAGATTGAGTGAAGCTTCGAGAGCGCTCGCGCGCT 983
Qy 198 GGAACAGATCTCGGTTCAGACCGCGAGAGACCGTCTGATTCGCGCGATGATGAGCC 257
Db 982 TGAACGATCTCGGTTCAGAGAGCGATATAGCGCGCGCGCGAGATCTCGCGCGCT 923
Qy 258 AGCGCGGAGCGCGGAGCGGCGGAGCGGCGGTCTGAGGTG-----CGCG 302
Db 922 AGCATGATCTCTGCGCGGTCTGAGCGGTCTGAGAGCGCGCGCTTGGCGAGCGCG 863
Qy 303 CGCGATGCGATCTCGGAGCTGCGCGGAGTGAAGCGAGATATCGGTGAGCGGAGTGA 362
Db 862 CGGTAGAGAGATGAGGAGCGGAGCGGAGCGGAGCGGAGATAGCGCGGTCTGAGTGA 803
Qy 363 ACGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 422
Db 802 CGAAGAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 746
Qy 423 TGGCGGTGATGAGCGCGGAGGAGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 482
Db 745 ACGAGTGGCGCGCGCGAGATGAGGAGAGCACCTGCTTCGCGCGAGATGATCGGCG 686
Qy 483 TCGAGCG 539
Db 685 TCTGCGAGCGCTTGGCGAGCGGTGAGCGCGCGAGTGGCGAGTGCAGACCTGCTGGG 626
Qy 540 GATCG 599
Db 625 ACTCG 566
Qy 600 GCGATGATGAGAGCGGTACCGGTCTGATCTGAGAGCGGTGCGCGCGCGCGCGCG 659
Db 565 GGTGATGAGAGAGGATGCGGTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 506
Qy 660 GCTCAAGCGCATGAGGCGCGGTGCGGTCTGATTCAGAGCGCGCGCGCGCGCGCG 719
Db 505 CATCAAGCGCGGTGCG 446
Qy 720 CGGTAGGTGCGGCGGTCTGAGCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 779
Db 445 GAGCGAGCG 386
Qy 780 GCGAGCG 830
Db 385 TCG 326
Qy 831 TATTCACCGCGAGTCAACCTTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 890
Db 325 GATCTCTCTCGGAAAGCGGTCTGAGCTTGGAGCGAGCGCGCGCGCGCGCGCGCG 266
Qy 891 CATGACCTGTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 950
Db 265 GGTAGCGCGGTCAAGAGCGAGTGGCATGTTGCGCGCGCGCGCGCGCGCGCGCGCG 206
Qy 951 CGCGCGCGAGATGAGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1010
Db 205 GATGACGAGCGCGGTTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 146
Qy 1011 AGACCGAGGTGCTGCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1070

```
Db      145 AGCTCAGGTGATGATTCGGGCGAGTCGAGGCTTCGATGATGACGACGCGCAT 86
Qy      1071 CCGGTCTGGGTGCGCAACGCGACGCGATGACGCGACGCGCGCGCGCGCA 1130
Db      85 CCGGCGGTGCGGTGCGGAGCCCGGACGCGATGACGAGCGCGCGCGCGCGATGA 26
Qy      1131 TCAGCAGATC 1140
Db      25 CGATGACGTC 16

RESULT 9
US-10-808-807-5/c
; Sequence 5, Application US/10808807
; Publication No. US20040253663A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Co., Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2365 US NA
; CURRENT APPLICATION NUMBER: US/10/808,807
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/477,874
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Pantoea agglomerans strain DC404
US-10-808-807-5

Query Match      19.9%; Score 228.6; DB 20; Length 1170;
Best Local Similarity 52.8%; Pred. No. 2e-48;
Matches 595; Conservative 0; Mismatches 514; Indels 18; Gaps 4;

Qy      29 TCGGGAGGCGACGCGCGCTGCGACACCGGAATGGGGCGCGCTCCGGTGAAGATGCA 88
Db      1136 TGGAGAGGCGCTTTGATCCCGCGACAAATCGGACCGCGCTTTCGCAAGCAGAGCGCG 1077
Qy      89 AGCCGGTCGCGCAATGTCAGGCGCGCGGATAGAAAGCGCTTCATGCGCGCTGCGCGCAG 148
Db      1076 GCGCGGTGCGCGGAGGAGGAGGCGCGGATAGAAAGCGCGCATGCTGCTGTCACG 1017
Qy      149 CCGTAGAACCGCTGACGACGCGGATAGCAAGCGTGGCGCGGCGACGCGCGAACAAGCATC 208
Db      1016 TGGTAAACCGCTGATCAAGCGCGCGCTGGTGGTGGTCCGCGCAGAAAGCATG 957
Qy      209 CGGTTCAGAGCGCGGAGAAAGCGGTGCGATCCGCGCATGATGGCGCCAGCGCGCAGC 268
Db      956 CGGTAAAGAGGCGGAAAAAGCGTGTGTCGCAATGCGCGCGCGGAATCGCGCATC 897
Qy      269 GCGCGACGCGCGGACCGCGTCTGTCAGAGTGCAGCGC---CGCATGGCATCCCGCATTCG 325
Db      896 AGCTGATAGAGCGCGCGCTGATGCACTCCCGGCGCTTGGCAATGGGCTCCCGCAGCGCA 837
Qy      326 GCGGCGTAGAGGCGACGAATATCCGATGACGCGGAGTGAACAAGCCCTGCGCGCGCAACC 385
Db      836 ACCGCGACGCGGACGAATAGCGGTGGTGGCATGGAACAGACCGCGCGCGCGCATG 777
Qy      386 GGCACCGCGCGCTGCGCGGTGTCGCGCGAAGACCTATGCGGTGATGGCGCGCGCGGATG 445
Db      776 ---ACGCGGTGATGATGAACTGGTGCAGAAAGCGCGCGCGATCGCGGACAGGGGTGATC 720
Qy      446 GCGAGGATCCCTCTTCGCGCGCATCTCTGCGC---GGTCAGCGCGCGCTGCGCGCA 502
Db      719 GCGAGGCGCGCGCTGCTTCACGCGACAGCGCGGAGAGTGTCCAGCGCTCTGCGCGGCG 660
Qy      503 TAGTCCAGGCGCGCTGCGCGCGCGCGCATGCTCCAGATCGCGCGCTGTAGCGCG 562
```

```
Db      659 TAATCGCATCCGCGCGCGGCTGAATCGCGCTGACCGCTCGGCGCGCTCAATGATGTC 600
Qy      553 GTATCTTCATGATGAGATGCGGATGGAATCTGAAGGCGACAGATAGATGAAGCGGATCCG 622
Db      599 GTGTCTTCATGATGAGGATGTGCGCGCTGAGCGCGGATGATGACAAACCGGTAGCGG 540
Qy      623 TCCATCTGCGGAACGGTGCCTGCATGATCATGCGGCGCTGACGCCCATGGGCGCGCTGCG 682
Db      539 TTGCGCTGGCGGACGCGGCGCATTCATCAGATCGGGCGCGTTAAACCGGTGGGCGCGGTC 480
Qy      683 GTCTGATCTGACGCGCCACGAATTTCTGGAACCCACGGTCAAGGTGGGGGTCTCCAGC 742
Db      479 AGTTCCCACTCTTCGCGCATGAAACGCTGATAGCCAAATGCTGAGGTGCGCGAGGCTGA 420
Qy      743 GCACACGCGGCGTGCATGACGAGCAGCCTCGATCCGCGAGCGCGTCCGTCAGCGTCGCG 802
Db      419 TAGCGCGGCGCATCATATCACGCGCGCGGATGAAGCGTCTGCTCGCTAGAGGTTACT 360
Qy      803 CCGGTATGCTCCAGCGTGCAGCATGCGTATTCACCGCGAGATGACACCTTCGACGAGC 862
Db      359 TCTTCGCGCGCTCACCGGTGACCGGTGACGTTTGTACGAATTTCTTTTCATGAGCGCG 300
Qy      863 CCGATCAGCGCGC-----CGCTCATTCAGCATAGCTTGTGTCAGCGCGCGC 913
Db      299 CGCATCGCTTGGGCAAAACGCGTGAAGGTGATGTCAGGTAGCGGTCTCATGACGAGTGGC 240
Qy      914 GAATGTCGGGAAACGCGCTCTGATCCGTTCATTCGCGCGCGACGAATGGGCGACAGG 973
Db      239 GACAGCTTCGAAAGATGACCTGTCACCGCTTCACGAGCATGGGCGACAGCGCGCGCAGC 180
Qy      974 CGCGCGACCATTCGCGGAAAGATCGGTGTGTGGAGAGACAGAGTGTGTCGAG 1033
Db      179 CAGCGGTCTGGGCTGGCGGTGATTCGTTGTGTAAGACAGAGTGTGTTCCGCGC 120
Qy      1034 GGGCGGACCGCGCTTCAGCATCACATGGCGCGATCCGCTTCGCGTCCGGAACGCGCA 1093
Db      119 GCGCGGTGCGCGCATCTCAGCATTAATACAGCAACGCTGAGATGACGCTGCTTATGTCG 60
Qy      1094 AGCGGATCGAGCGACCGGACGCGCGCGCGCGCGATCAGCAGATC 1140
Db      59 CAGCGATTAAGCCGCTTGGCGAGCGCGCGCGCGCATGATCAGATC 13

RESULT 10
US-10-997-844-6/c
; Sequence 6, Application US/10997844
; Publication No. US20050124033A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Pamela L
; APPLICANT: Bosak, Melissa
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; APPLICANT: Cheng, Qiong L
; TITLE OF INVENTION: Optimized Bacterial Host Strains of Methylomonas sp. 16A
; FILE REFERENCE: CL-2230 US NA
; CURRENT APPLICATION NUMBER: US/10/997,844
; PRIOR FILING DATE: 2004-11-24
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; PRIOR APPLICATION NUMBER: US 60/527,877
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Pantoea agglomerans
US-10-997-844-6

Query Match      19.9%; Score 228.6; DB 22; Length 5632;
Best Local Similarity 52.8%; Pred. No. 1.7e-48;
Matches 595; Conservative 0; Mismatches 514; Indels 18; Gaps 4;
```

OY	29	TCGGGCAAGCGACGGCCCTGCGACACGCGGAATGGCGGGGGCGTCCGGTACATACATCGA	88
Db	2014	TGAAACAGAGGCTTTGATGCGCCCGCAATCGGACCGGGCGGTTTGGCCAAAGACAGAGCGCG	1955
OY	89	AGCCGGTGGGCAATGTCAAGCGCCCGGCAATAGAAAGCCCTGATCAGCGCTGCGGAGG	148
Db	1954	GCGCGGTGGCGGAGCGAAAGCTGGCGGGCATTAAGACGGGGGATCAGTGTCTCGTAAGC	1895
OY	149	CGGTAGAAACGCTGCAAGGCGATGAGCAGGTGCGGCGGAGCGCGGACCGCGAAGCATC	208
Db	1894	TGGTAAACCGCTGATCAGCGCGCAGCGCTGATGCGGGTGTGCGGCGCAGGAAAGCATG	1835
OY	209	CGGTTCAAGACCGCGAGAGAGCGGTGCGGATTCGGCGCATGATAGGCCAGCGCGCAC	268
Db	1834	CGGTTAACACAGCGCGAAGAAACGTTGTGTCTGCGCATGAGCGCGCGCAATTCGCGATC	1775
OY	269	GCGGAGAGGGGAGAGCGCGGTGCTGAGATGCGGGCGC---CGGATGGCATTCGCGACCTGC	325
Db	1774	AGCTGATTAAGCGCGCCCTGATGCACTCCGCGCGGTTGGCAATGCGGTCCGCAAGCGA	1715
OY	326	GCGGATTAAGGAGCAATATCCGATGACGGGATGAGAACAGCCCTGCCCCAGGCCAAC	385
Db	1714	AACCGACAGGACAGGATTAAGCGGTGTGCGATAGAACACACCGCGCGCAGGCGCGCTG	1655
OY	386	GGCACCGCGCCCTGCGCGGTGTGCGCGCAAGCCTATGCGGTATGAGGCCAGCGCATG	445
Db	1654	---ACCGCTATGATGATGAACTGGGTGCGAAGGCGCGCGGATGCGCGGACAGGGTATC	1598
OY	446	GGCAGAGATGCCCTTTGCGCGCGCATCTCTGCCC---GCTCAGAGCCCGCGCTGGCGGCA	502
Db	1597	GGCAGCGCCCCCTGTTCTCTCAAGCACAGCGCGCGAGACTGCAAGCCCTGTGGCGGACG	1538
OY	503	TAGTCACAGCAGCGCTGCGCGCAGCGCGCATCTGTCAGATGCGCGCTGCTGTACGC	562
Db	1537	TAATGGCAATCCGCGCGGGGTGTAATCGGCGTGAAGCGTGGGCGGTCAATTAATGTC	1478
OY	563	GTATCTTCGATCAAGATCGGGGTGGGACTGAAGGGCAGCAGATAGATGAACCGGTACCG	622
Db	1477	GTGTCTTCGATTAAGAGAGGTGTGCGCGCTGAGCGGACAGGGATTAAGCAAGCGGTACCG	1418
OY	623	TCCATCTGCGGAACGCTGCGGTCCATATATCGGGGCTTGACGCCATGGGGGGCGTGC	682
Db	1417	TTGCCCCGTGGGAGCGCGGCGCATTCATCAAGATCGGGCGCGTTAAACCGGTGGGGCGCGTCT	1358
OY	683	GTCTCGATCTGACGCGCCAGCAATTTTCGAAACCCACCGTCAAGTGGGGGTCTCGACG	742
Db	1357	AGTTGCCACTCTTGCGCGATTAAGCCTTGATAGCCATAGCTGAAGGTGGGCGAGGGCTGA	1298
OY	743	GCACCAACGGGCGTGCATCAGCAGGACAGCCTCGATCCGCGAGCGCGTCTCAGCGTACG	802
Db	1297	TAGCGCGGGCATCAATACCGCGCCCGGGTAAAGGTGCGTCCGTGCTGTAAGGGTTACT	1238
OY	803	CCGGTATCTCAAGCGTGCACATGCGTATTCCACCGCAGATGACACCTTCGACAGC	862
Db	1237	TCTGCGCGCTCACCCGTGACAGCGTCACTGTTGTCAAGCAATTTCTTTTCATCAGCGCG	1178
OY	863	CCGATCAAGCGGCGC-----CGCTCGATTGACAGCATGAGCCTGTCTGTCAGCGCGCGC	913
Db	1177	CGCATCGCTTGGGCAAAACGCGTGAAGGTATGTGTCAGTATAGGCGGTACATGAGGGTGC	1118
OY	914	GAATGATGGGAAACGCAAGCTCTCGATCCGTCAATTCGCGCGACGAATGGGAGACAG	973
Db	1117	GACAGTTTCGAAAGTGAAGCTGTGTAACCGGTCCAGCGATGGGACCAAGCGGCGCGACG	1058
OY	974	GCGCCAGCCATTGGGAGAAAGATCCGTGTCTGTGGCAGACGACAGGTGTCTGTCCAG	1033
Db	1057	CAGCGGTCTGGGCTGGCGGTGATATCGTTGTGAAAGGACACAGATGTGGTTTCCGCG	998
OY	1034	GGCGCCGACCGCGGTGAGATACACATGGCGCATCCGGTCTGCGGTGCGCAACGGCA	1093
Db	997	GGCGGTGCGCGACTTCAGCATTAATACAGCAAGCGTGGATAGACGTGTGTTTATGTGCG	938

```

Dy      1094 AGCGCGATTCAGCGCACCGGACAGCGCCCGCGCGCGATTCAGCAATC 1140
Db      937 CAGCGCATTAAGCCCGCTTGCGCCAGCCCGCGCGACAGAAATCAGATC 891

RESULT 11
; US-10-808-807-18/c
; Sequence 18, Application US/10808807
; Publication No. US20040253663A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Co., Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2365 US NA
; CURRENT APPLICATION NUMBER: US/10/808, 807
; CURRENT FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/477,874
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 8814
; TYPE: DNA
; ORGANISM: Pantoea agglomerans strain DC404
; US-10-808-807-18

```

Query Match	19.9%	Score 228.6;	DB 20;	Length 8814;
Best Local Similarity	52.8%	Pred. No. 1.7e-48;		
Matches 595; Conservative	0;	Mismatches 514;	Indels 18;	Gaps 4;

OY	29	TTGGGCAAGGCAAGCCCAACGGCTCTGGGACAAGCGAATGAGGCGCGCCTCCGGTATGAGATGGA	88
Db	5559	TGAGACAGGGGCTTTTGATTCGCCCCCGACATCGGCACCGCGCTTTTCCACAGACGAGGGCGC	5500
OY	89	AGCGGCTCGGCAATGTCAGGCGCGCCGGCATATGAAGCGCTGATCAGCGCTTGGCGGAGG	148
Db	5499	GCGGGGTCGCGGAGCGAAGTGGCCGGGCATATAAAACGGGGGATCTACCTGCTCCTCAAGC	5440
OY	149	CGGTAGAACCGCTTGACAGAGGCGATAGCGACGTCGGGCGGGACCGCGGAACAGATC	208
Db	5439	TGTATAAACCGCTGCATCAGCGCCAGCGCTGGTCGGGTGTCGCGCACGGAATAAGCATG	5380
OY	209	CGGTTTCAGCAGACCGCAGAGAACGGTCCGCATCCGCGGATGATATGCTCCAGCGCGCAAC	268
Db	5379	CGGTTAAGCAGAGCGGAAAAAGCGTTGTGTCTGCGACATGGCGCGCCGGGAATTCGGCATTC	5320
OY	269	GCGGACGCGGCGAGACGCGGTCGTCAAGGTCGCGCGC---CGGATGGGATCCGCACTGTC	325
Db	5319	AGCTGATAGAGCGCGCCCTGATGCAATTCGCGCGCGCTTGGCAATCGGTCGCGCACGCCGA	5260
OY	326	GCGGCATAGGCGACGCAATATCCGGTACGCGGATGGAACAGCCTGCCCCAGGCCAAC	385
Db	5259	ACCGCCAGCGGCAACGATATGCGGTGTGGCATATGAACAGACCGGGCGCAGGCCGCTG	5200
OY	386	GGCACCGCCCCCTGCGGCTGCTGCGCGCAGAGCTTATGCGCTATAGGGCCAGCGCATG	445
Db	5199	---ACCGCTGATGATGAACTGGTTCAGAAAGCGCGCGGATGCGCGAACAGGGTATTC	5143
OY	446	GGCAGGATGCCCCCTTGGCGCGCGCATTCCTCGGCC---GGTCCAGCCCGGCTGGCGGCA	502
Db	5142	GCGACGCGCCCCCTGTCTCTCAACGACCAAGCGCGCAAGCTGCCAGGCCCTGTGGCGGACG	5083
OY	503	TAGTCAGCGACGCGCTCGCGCACGCGCCCATTCGTCCAGATCGCGCGCGCTGCTAGACG	562
Db	5082	TAAATGGCAATTCGCGCGCGGCTGAATTCGGCGGTGAGCGGTGGCGCGCTCAATGTATGTC	5023
OY	563	GTAATCTTCGATCAGGAATCGCGGTGGAGCTGAAGGCGACGATGATGAAGCGGTACCCG	622
Db	5022	GTCGCTTCGATTAAGCAGAGGTCTCGCGCCTAGACGCGCAGGATATAGACAAAGCGGTACCG	4963

Qy	623	TTCCATCTGCGGAAGAGGTGTCGTCCATATCATTCGGGGGCTCGAAGCCATGGGGGGCTCG	682
Db	4962	TTGCCCTTGAGGAGCGGGGCATCCATCAGATGAGAGGGGGCTTTAACCCGTGGGGGGCTC	4903
Qy	683	GTCCTGATCTTCGACGGCCACGAATTTCTGGAACCCACCGTCAGGTGCGGGGTCTCGAC	742
Db	4902	AGTTGGCACTCTTGCGCGATGAAAGCGCTGATAGCCATGTCTGAGGTGGGGGAAGGCTGA	4843
Qy	743	GCACCCAGGGGGGTGATATGACGAGCAGCCTTCGATTCGCCAGGCCGTCCGTCAAGCTCGG	802
Db	4842	TAGCCGGCGGCATCAATCACCGCCCGCGGTAAAGGTGCTGCTGCTGAGGGTTACT	4783
Qy	803	CCGGATTCGTCCAGGTGCGACATGCGTATTTCCACCGCAGATCGACACCCCTGAGCAGC	862
Db	4782	TCTGCGCCCTCAACCCCGTGAACGGGTACAGTTTGTCAAGAAATTTCTTTTCATAGCCCG	4723
Qy	863	CCGATCAGCGCGCC-----CGCTCGATCGAGCCATAGCCTGTGTCAGGGCGGC	913
Db	4722	CGCATCGCTTGAGGCAAAACGCGTGGAAGGTGATAGTCAAGTATGCGCTCATGCAAGGATCGC	4663
Qy	914	GAATGTGTGGAAACGCACTCTCTGATTCCTCCATTCGTCGCCGACCAATGGGGGAAG	973
Db	4662	GACACGTTCGAAAGTGAACGTCTGTACCCGTCACGCGATGGGCGCACCGAGGGCGCAGC	4603
Qy	974	CGCCCGACCCATTCGGGGGCGAAAGATTCGGTCTGTGCGACGACACAGTGTGCTGCCAG	1033
Db	4602	CAGCGTCTTGCGGCTGCGGTGATATCGTGTGTGTAAGAACACAGAGTGTGTTTCCGCG	4543
Qy	1034	GGGCCGACCGCGCGGTGAGCATCAAGATGCGCGCGATTCGGTCTGCGGTCCGAACGGCA	1093
Db	4542	GGCCCGTGGCGGCACCTCCAGATTAAATCAGCAAGCGTCGGATGACGCTGTTATGTGCG	4483
Qy	1094	AGCGGATCAGCGCACCGGACAGCCCGCGCGCGATCAGCAGATC	1140
Db	4482	CAGCGATAGACCCGTTTGCCAGCCCGCGCGACAGAAATCAATC	4436
RESULT 12			
US-10-997-844-41/C			
; Sequence 41, Application US/10997844			
; Publication No. US20050124033A1			
; GENERAL INFORMATION:			
; APPLICANT: Sharpe, Pamela L			
; APPLICANT: Bosak, Melissa			
; APPLICANT: Tao, Luan			
; APPLICANT: Sedkova, Natalia			
; APPLICANT: Cheng, Qiong L			
; TITLE OF INVENTION: Optimized Bacterial Host Strains of Methylomonas sp. 16A			
; FILE REFERENCE: CL-2230 US NA			
; CURRENT APPLICATION NUMBER: US/10/997,844			
; CURRENT FILING DATE: 2004-11-24			
; PRIOR APPLICATION NUMBER: US 60/527,083			
; PRIOR FILING DATE: 2003-12-03			
; PRIOR APPLICATION NUMBER: US 60/527,877			
; PRIOR FILING DATE: 2003-12-08			
; NUMBER OF SEQ ID NOS: 53			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 41			
; LENGTH: 8614			
; TYPE: DNA			
; ORGANISM: Pantoea agglomerans			
US-10-997-844-41			

Query Match	19.9%	Score 228.6	DB 22	Length 8814
Best Local Similarity	52.8%	Pred. No. 1,76-48		
Matches	595	Conservative 0	Mismatches 514	Indels 18
			Gaps	4

OY	29	TCGGCAGCAGCAGCGCCTTCGCACAGCGGAATGGCGCGGCGCTCCGCTGATGCATATGCA	88
Db	5559	TGGACACAGGCGCTTGATGATGCCGCCGACATCGGCACCGCGGCTTTGCCAAGCAGCAGGCGGC	55000
OY	89	AGCCGCGTGGCGCCATGTACGAGCGCCCGGACATAGAACCGCTGCATACGCGGCTGCGGCAGG	148

Db	5419	GCCTGGCTCGGCGGAGCGAAGCTGGCCGGCATTAATAACGGCGCATACGCTGTCGTCAAGC	54410
Oy	149	CGGTAGAACCCGCTGACAGCGCGATAGCGAGCTCGGCGGGCAGCGCGGAACAAGATC	208
Db	5439	TGGTAAACCGCTGCATCAAGCGCCAGCGCTGGTCGGGGTGTGCGCGCAGGAAAAAGCATG	5380
Oy	209	CGGTTTCAGCAGCCGACAGAAAGCGGTCCGCGATCCGCGCGATTCGATATGGCCACCCCGCAAC	268
Db	5379	CGGTAAAGCAGGCGGAAAAAGCGTTGTGTCTGCGACATGCGCGCGCGAATTCGAGCTTC	53220
Oy	269	GCGCGACGGGCGAGCGGCTGATCAGGTCGCGGCC---CGGATATGAGATCCGCGACCTTCG	325
Db	5319	AGCTGATAGAGCGGCGCTTGATGCAATCCGCGCGCGTTGGGAATGCGGTCCGCGACGCCA	5260
Oy	326	GCGGCATATAGGCGACCGAATATCCGCTGACGGGGTGGAAACAGCCCTTCCCAAGCCCAAC	385
Db	5259	ACCGCCAGCGGACGAAATAGCCGGTGGTGGCATGGAACAAGACCGCGCGCAGGCGCTG	52000
Oy	386	GGCACCGCGCCCTGCGGCTGGTCCGCGCAGAAAGCTATATGGGTTCATGGGCGCAGCGCGATG	445
Db	5199	---ACCGGCTGATATGAGAACTGGTGTGCGAAGGCGCGCGGATGCGCGAAGGGTATTC	51430
Oy	446	GGCAGAAATGCCCTTTTGCGCGCGCGGATCTCCGCGCC---GGTCCAGCCCGCGCTGGCGGCA	502
Db	5142	GCGACGCGCCCTTCTTCTCAAGCAACAGCGCGCGCAAGCTGCAGCCCTGCTGGCGGGCG	508
Oy	503	TAGTCCAGCGACGCTTGCGCCAGCGCGCGCATGTCAGATGCGCGCGCTGCTGTAGCGC	562
Db	5082	TAATCGGCAATCCGCGCGCGGGCTGAAATCGGCGTCTGAGCGCTCGGCGCGTCAATGATGTC	50220
Oy	563	GTATTCCTTCATCAGAGATGCGGGTGGGACTGAAGGGCAGCAGATATGATTAACGGGTACCG	622
Db	5022	GTGCTTCCTGATTAACAGAGGTGTGCGCGCTGAGCGGATTAAGACAAACCGGTATAGCGC	49660
Oy	623	TCCATCTGCGGAAAGGCTGCGCTGCATGATTCATCCGGGCGCTGAGCGCATGCGGCGGCTGC	682
Db	4862	TTGGCCCTGGGGAGACGGGGCATTCATCAGATTCGGGCGCGTTAAACCGTGGGGCGCGGTC	49020
Oy	683	GTCTCGATCTTCGACGCCCAAGAAATTTCTGGAACCCACGCGTCAAGTGCGGGTCCTCGACG	742
Db	4902	AGTTGCCACTCTGGCGGATGAAAGCGCTGATAGCAATGCTGAGGTCGGGAGGGCTGA	48420
Oy	743	GCACCAAGGGGCTGATTCACGACAGGAGCCCTCATCCGCGAGCGCGTCCGTCAAGGTCGCG	802
Db	4842	TAGCCGCGGCATATATCACCGCCCGCGGTAAGCCTGCTCGTCCCTGAAGGTTACT	47820
Oy	803	CCGGATTCGTCCAGCGTGCAGCATGTGCTATTCCACCGCAGATGCACACCTGACAGCAGC	862
Db	4782	TCTGCGCCGCTCACCCGTAACCGGTACAGTTTGTCAAGAAATTCCTTTTATACAGCCCG	47220
Oy	863	CCGATCAGCGCGCC-----CGCTCGATCGAGCCATAGCCTGTCTCAAGCGCGCGC	913
Db	4722	CGCATCGCTTGGGCAAAAACGCGTGAAGGTATGTCAGGTTAGCGGTCAATGACGGGTGCGC	46620
Oy	914	GAATGTCGGAAGAACGCACTCTCTGATTCGCTTCATTCGCGCGACAGTAATGGGCGACAGG	973
Db	4662	GACACGTTTCGAAAGTGAACGTCTGTAACCTGTCACGCGATAGGCGCAACGAGCGCGCCAGC	46020
Oy	974	CGCGCAGCGCATTTGGGGGGAAGAATCCGTCGCTGGGACAGACCAAGTGTGCTGTCGAG	10320
Db	4602	CAGGCGTCTGAGGCTGGGCGTATATCGTTTGGTGAAGACACAGGTGTGGTTTCGCGCG	45420
Oy	1034	GGGCGCGACCGCGGCTGAGCATCAAGATGCGCGCATCCGATCTGCGGTCCGGAACGGCA	10920
Db	4542	GGCGGTCGCGCGCACTCCAGCATTAATTCAGCAAGCGTCGGATGACGCTGCTTTTATGTCG	44820
Oy	1094	AGCGCGATCAGCGCACCGGAACAGCCCGCGCGCGCGCATCAACAGATC	1140
Db	4482	CAGCGATTAAGCCCGTTGTGGCCAGCCCGCGCGTACCAATATCAGATC	4436

RESULT 13


```

US-10-041-018-19/c
; Sequence 19, Application US/10041018
; Publication No. US20040072323A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seichi P.T.
; APPLICANT: Hart, Elizabeth A.
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism
; FILE REFERENCE: P02080US1/10025547
; CURRENT APPLICATION NUMBER: US/10/041,018
; PRIORITY FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259880
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 12753
; TYPE: DNA
; ORGANISM: Pantoea agglomerans
US-10-041-018-19

```

Query Match 18.6%; Score 213.4; DB 18; Length 12753;

Best Local Similarity 52.1%; Pred. No. 1.2e-44; Matches 581; Conservative 0; Mismatches 516; Indels 18; Gaps 4;

```

QY 41 CGCAGCGCCTGCGACAGCGGAATGGCGGGCGTCCGGTACATGCGAAGCCGGTCGGCC 100
DB 7916 CGCAGCGCCTTCCGCCAGCGGAACCGGTGGCTTCCGCTCAAAATCGGGCTTATCAAG 7857
QY 101 AATGTGAGGCGCCCGCGATAGAACGCTCGATCAGCGGCTGCGGCGAGCGGTAGAACCG 160
DB 7856 AGAGAGAGCGGACCGCGGTAAAGCGCTCTACGCTGGGCTCCGGCAGCCATAAAGCGC 7797
QY 161 TCGAGAGAGGATAGGAGCGGTGGCGCGGCGGACCGCGGAACGATCCGGTTCAGACG 220
DB 7796 TCGATACCGCGGACGCGGCTTCTCTCGCGCGCGCGGAGAAAGCATCCGGTTCAGACG 7737
QY 221 CGCAGAGAGC---GTCGCGATCCGCGCATCGATGAGCGCCAGCGCGGACCGCGGACG 277
DB 7736 CGGAGATTCCTGCTCGCGCGCGAGTGGCGGCTTCCGCAACTGCGGGTGAAGCTATGAGC 7677
QY 278 GCGGAGCGGCTGCTGAGTGGCGCGCGCGCGATGCGCATCCGCGACTGCGGCGATAGGCG 337
DB 7676 GGAAGCGTGGCGCGCGCGGCTGTGGCAATGGCGTGGCGGAGGCGCACCGCAGCGG 7617
QY 338 AGGAAATATTCGGTGAAGCGGCTGGAACAGCCCTGCGCGGACCGCAACCGGACCGCC 397
DB 7616 AGGAAATAGCATGTGTAGGCTGAATAGCCGCGGACCGGCTGCGGACCGGCGCGCC 7557
QY 398 TGGCGGTGGTGGCGGAGAGCCTATGGGCTCATGGGCGGAGCGCGGATGGGAGATGCC 457
DB 7556 GGGCATTCGCG---CCACAGAGCTGTGATGTACCGCGCGGATTAATCGGAGACAGCG 7500
QY 458 CTT---TCGCGCGCATCTCTGCGCGCGATCCCGCTGCGGCGGATAGTCCAGCGAC 514
DB 7499 GTCCTCTCGGTTCAAGCTGGGCGGAGCTGCGACCTTTGCTGAGACATAGTGTGTAAC 7440
QY 515 GCGTGGCGGAGCGCGCGCATGTCTCAATGCGCGCGCTGCTGAGCGGCTATCTTGATC 574
DB 7439 GTCGTGCGTGAAGCGCATTTATCATCAGCTGGGAGCATTTGGCGGTAGGCGGTATCTGATC 7380
QY 575 AGGATCGGGTGGAGCTGAAGGCGAGATAGATAGTAAGCGGATACCGGTCAATCTAGGAG 634
DB 7379 AGCAGGCTTCCGCGGAGAGCGGAGCGGTGTAGCAAAAGGATAGCCCTGCTGCTCGCC 7320
QY 635 AGCGTGGCTCATGATCATTCGGGCGCTCGACCGCATGGGAGCGGTGCTTCGATCTCG 694
DB 7319 ACCTGCGATCATTCAGATCGGTAGCGTACAGCGCGGTGGGCTGTGTACGCGCGCATCG 7260
QY 695 AGCGCCACGAATTTCTGGAACCCACAGGTTCAGGTGGGCGGTCTTCAGACGCGACCGAG 754
DB 7259 TGAACCAAGAAAGAGCTGATAGCGGTTCGATGCGGAACTGGCGGCTTCAGCGCGCTCG 7200
QY 755 TCGATCAGCGAGCGCATCTGATCCGCGAGCGGTCCGTGAGGTGCGCGCGGTATGCTCC 814

```

```

DB 7199 TGAATCAGCGCTCCGCGAGAGCGGCTCACCGTTGGCAAGGCGACGCTATGGGTAC 7140
QY 815 AGCGTGGCATATGCTGATTTCCACCGGATGCAACCC-----TCAGACAGCCCG 865
DB 7139 ACCTCGCTCACCGAAGAGTTTAGCCAGATGTTCTCCCGACGCGCTGATGCGGCGCTCG 7080
QY 866 ATAGAGCGCGCGCTCGATTCAGAGCATAGCGGTGTGTAGAGGCGGCGGAATGTCGGGA 925
DB 7079 GCAAGCGCTTGAAGTAAATGAGTATGAGTACCGCGCGGAGGCGGAGCGGAGATGGA 7020
QY 926 AACGCACTCTCTGATCCGTTCATTCGCGCGGACGATGAGGCGGAGCGCGGCGGACCAT 985
DB 7019 AACTGCACTCATAGCGCGCGGCGGAGGCGGCGGAGCGGCGGAGCGGCGGCGGCTCG 6960
QY 986 TCGGCGGAAGATCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1045
DB 6959 CCGGAGTGCAGATCGCTTCATGATGATGACAGGTATGCTCCCGCGGCGTGTCTCCCG 6900
QY 1046 GGTGAGATCAGATGAGTGGCGCATTCGCGGTGCGGCGGAGCGGAGCGGATCAGC 1105
DB 6899 GCTTCATCAGAGGAGGTTAAGCTGCGGTGAGCGCTGCGGAGCGGAGCGGATCAGC 6840
QY 1106 GCACCGAGAGCGCGCGCGCGCGCGGATCAGCAGATC 1140
DB 6839 CGGTTGCGAGGCGCGCGCGCGGATCAATCAATCAGATC 6805

```

RESULT 14

US-10-804-677-5/c

; Sequence 5, Application US/10804677

; Publication No. US20040224383A1

; GENERAL INFORMATION:

; APPLICANT: E.I. duPont de Nemours and Company, Inc.

; APPLICANT: Cheng, Qiong

; APPLICANT: Tao, Luan

; APPLICANT: Sedkova, Natalia

; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS

; FILE REFERENCE: CL2346 US NA

; CURRENT APPLICATION NUMBER: US/10/804,677

; PRIOR FILING DATE: 2004-03-19

; PRIOR APPLICATION NUMBER: US 60/468,596

; PRIOR FILING DATE: 2003-05-07

; PRIOR APPLICATION NUMBER: US 60/527,083

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 5

; LENGTH: 1167

; TYPE: DNA

; ORGANISM: Pectobacterium cyrripedii DC416

US-10-804-677-5

```

Query Match 18.5%; Score 213; DB 20; Length 1167;
Best Local Similarity 51.4%; Pred. No. 1.8e-44;
Matches 573; Conservative 0; Mismatches 530; Indels 12; Gaps 3;

```

```

QY 40 GCGCAGCGCTTGCAGACAGGGAATGGCGGCGGTGCGGATGCGAAGCGGCTGCGC 99
DB 1128 GCGCAGCGCTTGCAGAGAGGAGACCGGCGGTTGCGGACAGAAATGCGCTTTATCGCT 1069
QY 100 CATGTGAGCGCGCGCGCATAGAACGCTTCATCAGCGGCTGCGGAGCGGCTGTAAGACG 159
DB 1068 GAGCGGCAATTCGCGGCGGCTTAAACGCGTAAATCATTCGCGGTCAAGCGGTTAAAGG 1009
QY 160 TCGCAGAGGCGGATAGCGAGCGGTGCGGCGGAGCGCGGAGACAGATCGGTTAGCAG 219
DB 1008 CTGCAATGAGCGCGAGCGGCTGTTAGCGCGCTGCGGAGAAACAGATGCGATTTAAGC 949
QY 220 CCGCAGGAAGCGGTGCGCATTCGCGCGATGAGGCGGCGGCGCA--CGCGCGAGCG 277
DB 948 ACGGAAAGAGCGCTCTCGCGCGCATGTGACGGGCAAGGATGATGATGATGCTCAG 889

```

```
QY 278 GCGACGCGGTCGTCAAGTCGCGCGC-CGCGATGCGATCCGCGACCTGCGCGCATAGGG 336
DB CGTGTAGAGCTCGCGGGGCGAGCGTGTGCAATCTTCTCCGACGCGCACCGCGCGGG 829
QY 337 CAGCGCAATTCGCGGTGACGCGGGTGGAAACAGCCCTTCCCGCCCAACCGGACCGCCCC 396
DB 828 CAGCGCAATTCGCGGTGACGCGGGTGGAAACAGCCCTTCCCGCCCAACCGGACCGCCCC 769
QY 397 CTGCGCGGTGTCGCGCGCAGAAACCTATGAGCTATGAGGCGCAGCGGATGGGAGATGCC 456
DB 768 GTGCTGCTGTTGCCAGAACTGATTCGATTCGCGCGCTGAGCGGTATCGGACGAGACCGGTG 709
QY 457 CTTTTCGCGCGCGCATCTCTGCGCGGTCCAGCCCGCGCTGCGCGCATGATCCAGCGACGC 516
DB 708 CTCTCTGCGCGCAGCTGCGCGCAAAATTCAGCGCGCGGTGCGGTGCGCATATCCGTATGTG 649
QY 517 CTGCGCGCGCGCGCATCTGTCGAGATTCGCGCGCGGTGCGGTGCGGTATCCGTATGTG 576
DB 648 ACGCGCGCGCTGTGCGCGCATCCAGCGGTGATGATGATGATGATGATGATGATGATGATG 589
QY 577 GATGCGGCTGAGCTGAAAGGCGAGCAGATAGATGAACCGGTACCGGTCCATCTGCGGAAAC 636
DB 588 TAAAGCGCTGCGCATGAGCGGCGAGGCTGTAACAAAGCATTAACCTGCTGCTGATCGAC 529
QY 637 GGTGCGCTGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 696
DB 528 GCTGCGCTGATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 469
QY 697 GCGCGCAATTTCTGGAACCCAGCGTCAAGTTCGCGGGGTCTGCAAGCGCACCGCGCGTC 756
DB 468 ACCGCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 409
QY 757 GATGCGGCGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 816
DB 408 AATGACCGCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 349
QY 817 GGTGCGCATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867
DB 348 CTGCGTGAACGCGCGCTGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 289
QY 868 CAGCGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 927
DB 288 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 229
QY 928 CCGGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987
DB 228 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 169
QY 988 GCGCGCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1047
DB 168 TTGCGTGAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 109
QY 1048 GTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107
DB 108 CTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49
QY 1108 ACCGCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1142
DB 48 GTTACGCGCGCGCGCATCCAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 14
```

RESULT 15
US-10-804-677-18/c

Sequence 18, Application US/10804677

Publication No. US20040224383A1

GENERAL INFORMATION:

APPLICANT: E.I. duPont de Nemours and Company, Inc.

APPLICANT: Cheng, Qiong

APPLICANT: Tao, Luan

APPLICANT: Sedkova, Natalia

TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS

FILE REFERENCE: CL2346 US NA

CURRENT APPLICATION NUMBER: US/10/804,677

```
/ CURRENT FILING DATE: 2004-03-19
/ PRIOR APPLICATION NUMBER: US 60/468,596
/ PRIOR FILING DATE: 2003-05-07
/ PRIOR APPLICATION NUMBER: US 60/527,083
/ PRIOR FILING DATE: 2003-12-03
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO: 18
/ LENGTH: 8675
/ TYPE: DNA
/ ORGANISM: Pectobacterium cypripedii DC416
US-10-804-677-18

Query Match      18.5%; Score 213; DB 20; Length 8675;
Best Local Similarity 51.4%; Pred. No. 1.5e-44;
Matches 573; Conservative 0; Mismatches 530; Indels 12; Gaps 3;
```

```
QY 40 GCGACGCGCTGCGCAGCGGATGAGGCGGCGCTGCGGTGACGATGCGCAACCGGTGCGG 99
DB 453 GCGACGCGCTTCCGCGAGAGAGGACCGCGGTTTCCGCGACGAATGCGTCTTATGCT 4474
QY 100 CAATGTCAGGCGCGCGCATGAAAGCGCTCGATCAGCGGTGCGGACGCGGTGAAACCG 159
DB 4473 GAGCGCGATTTGCGCGGCTTAAAGCGCTTATCAATCCGCGTCAAGCCGTTAAACG 4414
QY 160 CTGACGAGCGGATGACGACGCTGCGGCGGCGAGCGCGGAAACGATCCGTTTCAAG 219
DB 4413 CTGATTTACGCGCGCGCGCTGTTCAAGCGCGTCCGCGAAGAAACGATGCTTTTACG 4354
QY 220 CCGCAGAAAGCGGTGCGATCCGCGCATCGGCGATGAGCGCGCGCA--CGCGGACGG 277
DB 4353 ACGGAAAAAGCGCTGCTGCGCGCGCATGCTGACGCGGCAAAAGATTCGATGAGTGGCTCAG 4294
QY 278 GCGACGCGGCTGCTGAGTTCGCGCGC--CGCGATGCGATCCGCGACCTGCGCGCATAGG 336
DB 4293 GGTGAGCGCTGCGCGCGCGAGCGTGTGCAATCTTCTCCGCGCGCGCACCGCGCGG 4234
QY 337 CAGCGCAATTCGCGTGAACGAGGATGAAACAGCGCTGCGCGCGCGCAACCGCGCACCGCGCC 396
DB 4233 CAGCGATGAAACGCGTGTGCGGTGAAACAGTCCGCGCGCGCGCGCGCGTTCGCG 4174
QY 397 CTGCGCGGTGTCGCGCGCGCAAGACCTTATGCGCTCATGCGCGCGCGCGAGATGCC 456
DB 4173 GTGCTGCTGTTGCGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 4114
QY 457 CTTTTCGCGCGCGCATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 516
DB 4113 CTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4054
QY 517 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 576
DB 4053 ACGCGCGCGCGTGTGCGCGCATCGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3994
QY 577 GATGCGGCTGGAATGAAAGCGCGCATGATGATGATGATGATGATGATGATGATGATGATG 636
DB 3993 TAAAGCGCTGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3934
QY 637 GGTGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 696
DB 3933 GGTGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3874
QY 697 GCGCGCAATTTCTGGAACCCAGCGTCAAGGTGCGGCGTTCGACGCGCGCAACCGGCGGTC 756
DB 3873 ACGCGCAATGCTGATGAGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3814
QY 757 GATCAGCGAGCGCGTTCGATCCGCGAGCGCGTCCGTCAGCGGTGCGCGCGGTATGCTCA 816
DB 3813 AATCAGCGCGCGCGTTCGATCCGCGAGCGCGTCCGTCAGCGGTGCGCGCGGTATGCTCA 3754
QY 817 GGTGCGCGCATGCGATTCACCGCGAGTTCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 867
DB 3753 CTGCGTACCGCGCGCGTGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 3694
```

```
OY 868 CAGCGCGCCGCTCGATCGAGCCATAGCCTGTCTCAGGCGCGCGAATGTCGGGAA 927
Db 3693 AAAATCTCTCGAGGCAATCGAGCAATATTCACCATCCAGCTGGCGAGCGAGGTGGGAA 3634
OY 928 CCGGACCTCTCTGATCCGTCATTTCGCCGACGAATGGCGACAGGCGCGCCAGCCATTTC 987
Db 3633 GCGTACCTGATTAACCGGGCCAGCGCGCCGAAAGCAGCGGCTCGAGCCAGCGAAACTGCGC 3574
OY 988 GGGCGAAAGATCCGTGTCTGTGCGAGGACAGAGTGTCTGTCCGAGGGGCCGACCGCGC 1047
Db 3573 TTGCTGACGCTCTTCGCGATGGAACGACGAGTATGATTCGCGCGCTGCGCTGACT 3514
OY 1048 GTCGAGCATCAAGATCGCGCATCCGCTCTGCGGTGCGAACGCGCAAGCGCGATCAGCGC 1107
Db 3513 CTCGAGTAGCAAAACCTTAAGTCGGGCTGCACTGGCGTAAACGACGCGCAATCAGCCC 3454
OY 1108 ACCGACAGCCCCCGCGCGCGCGATCGACGATCAT 1142
Db 3453 GTTAGCCAGGCGCGCACGACGACGATGACATCAT 3419
```

Search completed: June 22, 2005, 04:23:24
Job time : 832.57 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 19:57:05 ; Search time 309.581 Seconds
(without alignments)
6072.980 Million cell updates/sec

Title: US-10-695-980-1_COPY_5794_6942

Perfect score: 1149
Sequence: 1 tcatgctctctctctgcagca.....atcagcagatcatgctcat 1149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6CTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfillseq1.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1149	100.0	1149	3 US-08-660-645A-8	Sequence 8, Appl1
C 2	1149	100.0	1149	3 US-09-298-718-8	Sequence 8, Appl1
C 3	1149	100.0	1149	3 US-09-546-969-8	Sequence 8, Appl1
C 4	1149	100.0	1149	3 US-09-547-267-8	Sequence 8, Appl1
5	1149	100.0	8625	3 US-08-980-832-1	Sequence 1, Appl1
6	1149	100.0	8625	4 US-09-920-923B-1	Sequence 1, Appl1
7	1149	100.0	11233	3 US-08-980-832-27	Sequence 27, Appl1
8	1149	100.0	11233	4 US-09-920-923B-27	Sequence 27, Appl1
C 9	640.8	55.8	1161	1 US-08-663-310-5	Sequence 5, Appl1
C 10	640.8	55.8	1161	2 US-09-006-491-5	Sequence 5, Appl1
C 11	640.8	55.8	1161	3 US-09-335-919-5	Sequence 5, Appl1
C 12	640.8	55.8	2886	1 US-08-663-310-7	Sequence 7, Appl1
C 13	640.8	55.8	2886	2 US-09-006-491-7	Sequence 7, Appl1
C 14	640.8	55.8	2886	3 US-09-335-919-7	Sequence 7, Appl1
C 15	200	17.4	1235	1 US-08-095-726-13	Sequence 13, Appl1
C 16	200	17.4	1235	1 US-08-096-623A-13	Sequence 13, Appl1
C 17	198.4	17.3	1235	1 US-08-095-726-15	Sequence 15, Appl1
C 18	198.4	17.3	1235	1 US-08-096-623A-15	Sequence 15, Appl1
C 19	174.8	15.2	1149	1 US-07-783-705A-9	Sequence 9, Appl1
C 20	174.8	15.2	6918	1 US-07-783-705A-13	Sequence 13, Appl1
C 21	138.8	12.1	1631	1 US-08-663-310-12	Sequence 12, Appl1
C 22	138.8	12.1	1631	2 US-09-006-491-12	Sequence 12, Appl1
C 23	138.8	12.1	1631	3 US-09-335-919-12	Sequence 12, Appl1
C 24	59.2	5.2	1998	3 US-09-382-106-1	Sequence 1, Appl1
C 25	59.2	5.2	2229	4 US-09-252-991A-3240	Sequence 3240, Ap
C 26	59.2	5.2	2709	4 US-09-252-991A-3326	Sequence 3326, Ap
C 27	56.2	4.9	1254	4 US-09-252-991A-11478	Sequence 11478, A

C 28	56.2	4.9	1524	4 US-09-252-991A-11455	Sequence 11455, A
C 29	56.2	4.9	2028	4 US-09-252-991A-11416	Sequence 11416, A
C 30	55.8	4.9	1125	4 US-09-252-991A-6994	Sequence 6994, Ap
C 31	55.8	4.9	1929	4 US-09-252-991A-7017	Sequence 7017, Ap
C 32	55.8	4.9	1947	4 US-09-252-991A-7111	Sequence 7111, Ap
C 33	54.4	4.7	2364	4 US-09-902-540-6412	Sequence 6412, Ap
C 34	54.4	4.7	2461	4 US-09-902-540-6412	Sequence 434, App
C 35	53.8	4.7	4257	2 US-08-690-473-1	Sequence 1, Appl1
C 36	53.8	4.7	4257	3 US-08-823-288A-1	Sequence 1, Appl1
C 37	53.8	4.7	4257	3 US-08-843-658-1	Sequence 1, Appl1
C 38	53.8	4.7	4257	4 US-09-825-288A-1	Sequence 1, Appl1
C 39	53.8	4.7	12001	1 US-08-458-568A-11	Sequence 11, Appl1
C 40	52.8	4.6	12173	4 US-09-902-540-1022	Sequence 1022, Ap
C 41	52.4	4.6	741	4 US-09-724-797-73	Sequence 73, Appl1
C 42	52.4	4.6	3783	4 US-09-902-540-4846	Sequence 4846, Ap
C 43	52.4	4.6	21758	4 US-09-902-540-1238	Sequence 1238, Ap
C 44	52	4.5	4826	3 US-09-192-983-3	Sequence 3, Appl1
C 45	51.8	4.5	1294	4 US-10-151-832-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-660-645A-8/c

Sequence 8, Application US/08660645A

Patent No. 6087152

GENERAL INFORMATION:

APPLICANT: Hohmann, Hans-Peter

APPLICANT: Pasamontes, Luis

APPLICANT: Tessier, Michel

APPLICANT: Van Loon, Adolphus

TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann-La Roche Inc.

STREET: 340 Kingsland Street

CITY: Nutley

STATE: NJ

COUNTRY: USA

ZIP: 07110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/660,645A

FILING DATE: 07-JUN-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 95108888.9

FILING DATE: 09-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Pokras, Bruce A.

REGISTRATION NUMBER: 32,748

REFERENCE/DOCKET NUMBER: RAN 6002/170

TELECOMMUNICATION INFORMATION:

TELEPHONE: (201) 235-5801

TELEFAX: (201) 235-2363

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1149 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-660-645A-8

Query Match

Best Local Similarity 100.0%; Score 1149; DB 3; Length 1149;

Matches 1149; Conservative 0; Pred. No. 1,1e-225; Indels 0; Gaps 0;

QY 1 TCATGCTCTCTCTGAGCAGAGGGGGGCTTCGGGCAAGCAGCCGCTGCGACAGCGG 60
DB 1149 TCATGCTCTCTCTCTGAGCAGAGGGGGGCTTCGGGCAAGCAGCCGCTGCGACAGCGG 1090
QY 61 AATGGGCGGGCGCTCCGGTGAATGCGAAGCCGGTGGCAATGTATAGAGCGCCCGGCATA 120
DB 1089 AATGGGCGGGCGCTCCGGTGAATGCGAAGCCGGTGGCAATGTATAGAGCGCCCGGCATA 1030
QY 121 GAAGCGCTGATGATGAGGCGCTGCGGAGGCGGTAGAAACCGCTGAGAGGCGATAGCGAGC 180
DB 1029 GAAGCGCTGATGATGAGGCGCTGCGGAGGCGGTAGAAACCGCTGAGAGGCGATAGCGAGC 970
QY 181 GTCGGGCGGGCAGCCGCGAAGCAGATCCGGTTACAGACCCGAGAGACGGTCCGATC 240
DB 969 GTCGGGCGGGCAGCCGCGAAGCAGATCCGGTTACAGACCCGAGAGACGGTCCGATC 910
QY 241 CGCGCGATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 909 CGCGCGATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 850
QY 301 CGCGCGATGATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 849 CGCGCGATGATGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 790
QY 361 GAAACAGCCCTGCGCCCGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 420
DB 789 GAAACAGCCCTGCGCCCGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 730
QY 421 TATGGGCTCATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 729 TATGGGCTCATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 670
QY 481 GGTCCAGCCCGCGCTGGCGGCGATGTCAGCGAGCGCTGCGGCGGCGGCGGCGGCGGCGG 540
DB 669 GGTCCAGCCCGCGCTGGCGGCGATGTCAGCGAGCGCTGCGGCGGCGGCGGCGGCGGCGG 610
QY 541 ATCGCGCGCGTCTGCTGAGCGCGGTATCTCGATCAGAGATGCGGGTGGAACTGAAAGGCG 600
DB 609 ATCGCGCGCGTCTGCTGAGCGCGGTATCTCGATCAGAGATGCGGGTGGAACTGAAAGGCG 550
QY 601 CAGATGATGAAAGCGGTATCCGCTCATTCGCGGAAAGCGTCCGCTCATGATCATTCGGGG 660
DB 549 CAGATGATGAAAGCGGTATCCGCTCATTCGCGGAAAGCGTCCGCTCATGATCATTCGGGG 490
QY 661 CTCGACGCGCATGGGGGCGCTCGGTCTCGATCTCGATCTCGACCGCAAGATTTTGTGAAACCC 720
DB 489 CTCGACGCGCATGGGGGCGCTCGGTCTCGATCTCGATCTCGACCGCAAGATTTTGTGAAACCC 430
QY 721 GGTCAAGGTGCGGGGTCTCGACCGGCGACACCGGCGTCAATCAGCAGGAGCGCTCGATCCG 780
DB 429 GGTCAAGGTGCGGGGTCTCGACCGGCGACACCGGCGTCAATCAGCAGGAGCGCTCGATCCG 370
QY 781 CGAGCGGTCTCGATGAGGTGCGGCGGCGGTATCTCGATCTCGACCGCAATGCGTATTCACCG 840
DB 369 CGAGCGGTCTCGATGAGGTGCGGCGGCGGTATCTCGATCTCGACCGCAATGCGTATTCACCG 310
QY 841 CAGATGACACCTCTGAGAGGCGCCGATCAGCGGCGCGCTCGATGAGAGCGGCTCTG 900
DB 309 CAGATGACACCTCTGAGAGGCGCCGATCAGCGGCGCGCTCGATGAGAGCGGCTCTG 250
QY 901 CGTCAAGGCGGGCGGAAATGCTCGGGAACCGGACCTTCGATCCGCTCATTCGCGCGGAGC 960
DB 249 CGTCAAGGCGGGCGGAAATGCTCGGGAACCGGACCTTCGATCCGCTCATTCGCGCGGAGC 190
QY 961 AATGGGCGGAGCG 1020
DB 189 AATGGGCGGAGCG 130
QY 1021 GTGCTGTCTCGAGGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
DB 129 GTGCTGTCTCGAGGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 70
QY 1081 GTGCGGAACGGCAAGCGCGATCAGCGCACCGGACAGCCCGCGCGCGCGGATCAGCAGATC 1140

DB 69 GTGCGGAACGGCAAGCGCGGATCAGCGCACCGGACGCCCGCGCGCGGATCAGCAGATC 10
QY 1141 ATGGCTCAT 1149
DB 9 ATGGCTCAT 1
RESULT 2
US-09-298-718-8/c
; Sequence 8, Application US/09298718
; Patent No. 6124113
; GENERAL INFORMATION:
; APPLICANT: Hohmann, Hans-Peter
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tesler, Michel
; APPLICANT: van Loon, Adolphus
; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingstland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/298,718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/660,645
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pokras, Bruce A.
; REGISTRATION NUMBER: 32,748
; REFERENCE/DOCKET NUMBER: RAN 6002/170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-5801
; TELEFAX: (201) 235-2363
; INFORMATION FOR SEO ID NO.: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-298-718-8
Query Match 100.0%; Score 1149; DB 3; Length 1149;
Best Local Similarity 100.0%; Pred. No. 1.1e-225;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGCTCTCTCTCTGAGCAGAGGGGGGCTTCGGGCAAGCAGCCGCTGCGACAGCGG 60
DB 1149 TCATGCTCTCTCTCTGAGCAGAGGGGGGCTTCGGGCAAGCAGCCGCTGCGACAGCGG 1090
QY 61 AATGGGCGGGCGCTCCGGTGAATGCGAAGCCGGTGGCAATGTATAGAGCGCCCGGCATA 120
DB 1089 AATGGGCGGGCGCTCCGGTGAATGCGAAGCCGGTGGCAATGTATAGAGCGCCCGGCATA 1030
QY 121 GAAGCGCTGATGATGAGGCGCTGCGGAGGCGGTAGAAACCGCTGAGAGGCGATAGCGAGC 180
DB 1029 GAAGCGCTGATGATGAGGCGCTGCGGAGGCGGTAGAAACCGCTGAGAGGCGATAGCGAGC 970
QY 181 GTCGGGCGGGCAGCCGCGAAGCAGATCCGGTTACAGACCCGAGAGACGGTCCGATC 240
DB 969 GTCGGGCGGGCAGCCGCGAAGCAGATCCGGTTACAGACCCGAGAGACGGTCCGATC 910

Oy	241	CGCGGAGATCGAATGAGCCGCAAGCCGCGACCGCGCGGAGCGGGGCGAGCGCGGTGTGTCAGAGTCGCG	300
Db	909	CGCGGAGATCGAATGAGCCGCAAGCCGCGACCGCGCGGAGCGGGGCGAGCGCGGTGTGTCAGAGTCGCG	850
Oy	301	CGCGCGGATGGCATTCGCGGACCTTCGCGCGGATAGAGGACAGCGAATATTCGCGTGAACGAGGTG	360
Db	849	CGCGCGGATGGCATTCGCGGACCTTCGCGCGGATAGAGGACAGCGAATATTCGCGTGAACGAGGTG	790
Oy	361	GAAACAGCCCTTCGCGCGGACCGCGGACCGCGCGGCGCGGTGTGCGCGCAGAGGCC	420
Db	789	GAAACAGCCCTTCGCGCGGACCGCGGACCGCGCGGCGCGGTGTGCGCGCAGAGGCC	730
Oy	421	TATGGCGGTCAATGGGCGCAGCGGAGATGGGCGAGATGGCCCTTTGGCGCGCGATCTCTCGGCC	480
Db	729	TATGGCGGTCAATGGGCGCAGCGGAGATGGGCGAGATGGCCCTTTGGCGCGCGATCTCTCGGCC	670
Oy	481	GATCCAGGCCCGCTTCGCGCGGACATAGTCCAGCGACGCTTCGCGCAGCGCGCCATCTGTCAG	540
Db	669	GATCCAGGCCCGCTTCGCGCGGACATAGTCCAGCGACGCTTCGCGCAGCGCGCCATCTGTCAG	610
Oy	541	ATCGCGCGCGTTCGCTGTAGCGCGGTATCTTCGATTCAGAGATGCGGGTGGGACTGAAAGGCGAG	600
Db	609	ATCGCGCGCGTTCGCTGTAGCGCGGTATCTTCGATTCAGAGATGCGGGTGGGACTGAAAGGCGAG	550
Oy	601	CAGATAGATGGAAGCGGATACCGGTCCATCTGCGGGGAAAGCGTGGGTCCATGATATATCGGGCG	660
Db	549	CAGATAGATGGAAGCGGATACCGGTCCATCTGCGGGGAAAGCGTGGGTCCATGATATATCGGGCG	490
Oy	661	CTCGACGCGCATAGGAGGGGCGTCGGTCTCGATCTCGACGCGCCACGAATTTCTGGAACCCGAC	720
Db	489	CTCGACGCGCATAGGAGGGGCGTCGGTCTCGATCTCGACGCGCCACGAATTTCTGGAACCCGAC	430
Oy	721	GATCAGGTGCGGGGATCTCGACGCGCACGACGGCGTTCGATCAGCGCAGCGACGCTTCGATCGG	780
Db	429	GATCAGGTGCGGGGATCTCGACGCGCACGACGGCGTTCGATCAGCGCAGCGACGCTTCGATCGG	370
Oy	781	CGAGCGCGTTCGTCAGCGGTTCGCGCGCGGTATCTTCGACAGCGTTCGCGACATGCTATTTCCACCG	840
Db	369	CGAGCGCGTTCGTCAGCGGTTCGCGCGCGGTATCTTCGACAGCGTTCGCGACATGCTATTTCCACCG	310
Oy	841	CAGATTCGACACCTTCGACGAGCGCCGATCAGCGCGCGCGCTTCGATTCGAGCCATAGCTGT	900
Db	309	CAGATTCGACACCTTCGACGAGCGCCGATCAGCGCGCGCGCTTCGATTCGAGCCATAGCTGT	250
Oy	901	CGTCAGCGCGCGCGGATGCTCGGGGAAAGCGGACCTTCGATCCGTCATTTCGCGCGGACG	960
Db	249	CGTCAGCGCGCGCGGATGCTCGGGGAAAGCGGACCTTCGATCCGTCATTTCGCGCGGACG	190
Oy	961	AATGGGGCGACAGGCGCGCGCCGACGTCATTTCGCGGCGAAAGATTCCTGTGTGTGGCAGAACGAGT	1020
Db	189	AATGGGGCGACAGGCGCGCGCCGACGTCATTTCGCGGCGAAAGATTCCTGTGTGTGGCAGAACGAGT	130
Oy	1021	GTCGCTGTCGCGAGGGGCGGAGCCGCGCGGTGCGAGCATCAAGATGCGCGCATCCGGTCTGGC	1080
Db	129	GTCGCTGTCGCGAGGGGCGGAGCCGCGCGGTGCGAGCATCAAGATGCGCGCATCCGGTCTGGC	70
Oy	1081	GTCGCGAACGCGCAAGCGCGATTCAGCGCAACCGGACAGCCCGCGCGCGATTCAGGACATC	1140
Db	69	GTCGCGAACGCGCAAGCGCGATTCAGCGCAACCGGACAGCCCGCGCGCGATTCAGGACATC	10
Oy	1141	ATGGCTCAT 1149	
Db	9	ATGGCTCAT 1	

RESULT3
US-09-546-969-8/c
; Sequence 8, Application US/0954696S
; Patent No. 6207409
; GENERAL INFORMATION:
; APPLICANT: Hohmann, Hans-Peter
; APPLICANT: Paasmontes, Luis

```

APPLICANT: Tessier, Michel
APPLICANT: van Loon, Adolphus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/546,969
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: RAN 6002/170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-546-969-8

Query Match 100.0%; Score 1149; DB 3; Length 1149;
Best Local Similarity 100.0%; Pred. No. 1,1e-225;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGCTCTCTCTGTGACGAGGGGCGTTCCGGCAGGCGACGCGACTTGTGACACGG 60
Db 1149 TCATGCTCTCTCTCTGTGACGAGGGGCGTTCCGGCAGGCGACGCGACTTGTGACACGG 1090

QY 61 AATGGGCGGGCGTTCGGGTGACGATGCCAAGCCGGTCCGCCAATGTCAAGGCGCCCGGCATA 120
Db 1089 AATGGGCGGGCGTTCGGGTGACGATGCCAAGCCGGTCCGCCAATGTCAAGGCGCCCGGCATA 1030

QY 121 GAAGCGCTGCATCAGCGCGCTGCGGCGAGCGGTGAAACCGCTGACGAGCGGATGCGACG 180
Db 1029 GAAGCGCTGCATCAGCGCGCTGCGGCGAGCGGTGAAACCGCTGACGAGCGGATGCGACG 970

QY 181 GTGGGCGGGCAGCGCGGGAACAGCATCCGTTTCAGCAGCCGCGAGAAAGCGTTCGCGATC 240
Db 969 GTGGGCGGGCAGCGCGGGAACAGCATCCGTTTCAGCAGCCGCGAGAAAGCGTTCGCGATC 910

QY 241 CGGCGCATGCATGGCCCAAGCCGGGCATCCGGCGACCGGCGGACCGCGGTCTGCAAGTGGCG 300
Db 909 CGGCGCATGCATGGCCCAAGCCGGGCATCCGGCGACCGGCGGACCGCGGTCTGCAAGTGGCG 850

QY 301 CGCGCGCATGGCATCCGCGCATCCGCGCGGATGAGGGGAGGAAATATCCGGTGAACGGGGTG 360
Db 849 CGCGCGCATGGCATCCGCGCATCCGCGCGGATGAGGGGAGGAAATATCCGGTGAACGGGGTG 790

QY 361 GAACAGGCTTGCCGCCCAAGCCGAACCGGACCGGCGCCCTCTGCGGTGTGCGCGCAAGAGCC 420
Db 789 GAACAGGCTTGCCGCCCAAGCCGAACCGGACCGGCGCCCTCTGCGGTGTGCGCGCAAGAGCC 730

QY 421 TATGGCGTCAATGGGCGACGCGCATGGGCGAGATGCCCTTTCCGCGCGCATCTTCGCGCC 480

```

Db 729 TATGGGTCATGAGGCGAGCGATGAGGAGATGCCCTTTTCGCGCGCATCTCTGCC 670
Qy 481 GGTCCAGCCCCGCTTGCGGCGATAGTCAGCGACGCTTGCGCGAGCGGCATCTGCGAG 540
Db 669 GGTCCAGCCCCGCTTGCGGCGATAGTCAGCGACGCTTGCGCGAGCGGCATCTGCGAG 610
Qy 541 ATCGCCGCGCTGCTGTAGCGCGTATCTCGATCAGATGCGGGTGGAGCTGAAGGGCAG 600
Db 609 ATCGCCGCGCTGCTGTAGCGCGTATCTCGATCAGATGCGGGTGGAGCTGAAGGGCAG 550
Qy 601 CAGATAGATGAAGCGGTACCGCTCCATCTGCGGAAAGCGTGCATCATATCATGCGGGG 660
Db 549 CAGATAGATGAAGCGGTACCGCTCCATCTGCGGAAAGCGTGCATCATATCATGCGGGG 490
Qy 661 CTCGACGCCATGAGGGGCGCTCGGCTCGATCTCGATCGAGCCAGAAATTTCTGAAACCAC 720
Db 489 CTCGACGCCATGAGGGGCGCTCGGCTCGATCTCGATCGAGCCAGAAATTTCTGAAACCAC 430
Qy 721 GGTCAAGTCCGAGGCTTCGACGCGACACGCGGCGTTCATCAGCGAGCGCTCGATCCG 780
Db 429 GGTCAAGTCCGAGGCTTCGACGCGACACGCGGCGTTCATCAGCGAGCGCTCGATCCG 370
Qy 781 CGAGCGCTCCGTCAGGTCGCGCGCGGTCATCTGTCAGCGCGGTCGAGTATGCTATCCACG 840
Db 369 CGAGCGCTCCGTCAGGTCGCGCGCGGTCATCTGTCAGCGCGGTCGAGTATGCTATCCACG 310
Qy 841 CAGATCGACCCCTGACGAGCGCCGATCAGCGCGCGCGCTCGATCGAGCCATAGCCTGT 900
Db 309 CAGATCGACCCCTGACGAGCGCCGATCAGCGCGCGCGCTCGATCGAGCCATAGCCTGT 250
Qy 901 CGTCAGCGCGCGGAAATGCTCGGAAAGCGACCTTCGATCTGTCATCTGCGCGGACG 960
Db 249 CGTCAGCGCGCGGAAATGCTCGGAAAGCGACCTTCGATCTGTCATCTGCGCGGACG 190
Qy 961 AATGGCGGACAGCGCGCGCGCATCTCGCGCGGAAAGATCCGTCGTCGCAAGACAGAGT 1020
Db 189 AATGGCGGACAGCGCGCGCGCATCTCGCGCGGAAAGATCCGTCGTCGCAAGACAGAGT 130
Qy 1021 GTGCTGTCCGAGGGCGCGAGCCGCGCTCGAGCATCAGATGCGCGCATCTGCTGCG 1080
Db 129 GTGCTGTCCGAGGGCGCGAGCCGCGCTCGAGCATCAGATGCGCGCATCTGCTGCG 70
Qy 1081 GTGCGCAACGCGCAAGCGCATCAGCGCACCGGACCGCGCGCGCGCGCATCAGAGATC 1140
Db 69 GTGCGCAACGCGCAAGCGCATCAGCGCACCGGACCGCGCGCGCGCATCAGAGATC 10
Qy 1141 ATGGCTCAT 1149
Db 9 ATGGCTCAT 1
RESULT 4
US-09-547-267-8/c
Sequence 8, Application US/09547267
Patent No. 6613543
GENERAL INFORMATION:
APPLICANT: Hohmann, Hans-Peter
APPLICANT: Pasamonte, Luis
APPLICANT: Tessier, Michel
APPLICANT: van Loon, Adolpinus
TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/547,267
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/660,645
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pokras, Bruce A.
REGISTRATION NUMBER: 32,748
REFERENCE/DOCKET NUMBER: PAM 6002/170
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-5801
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-547-267-8
Query Match 100.0%; Score 1149; DB 4; Length 1149;
Best Local Similarity 100.0%; Pred. No. 1,1e-225;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCATGCTCTCTCTGTCAGAGGCGGCTTCGCGGCGAGGCGACGCGCTGCGACAGCGG 60
Db 1149 TCATGCTCTCTCTGTCAGAGGCGGCTTCGCGGCGAGGCGACGCGCTGCGACAGCGG 1090
Qy 61 AATGGCGGCGGCTCCGCTGACGATGCGAAGCGGTCCGCAATGTTCAGCGCGCGGCATTA 120
Db 1089 AATGGCGGCGGCTCCGCTGACGATGCGAAGCGGTCCGCAATGTTCAGCGCGCGGCATTA 1030
Qy 121 GAAGCGCTCGATCAGCGCGCTGCGGCGGCTAGAACCGCTGTCAGACAGCGCATGACGACG 180
Db 1029 GAAGCGCTCGATCAGCGCGCTGCGGCGGCTAGAACCGCTGTCAGACAGCGCATGACGACG 970
Qy 181 GTGCGGCGGCGGCGGCGGAAACGATCGGTTTCAGCGCGCGGAAAGCGGTCGCGATC 240
Db 969 GTGCGGCGGCGGCGGCGGAAACGATCGGTTTCAGCGCGCGGAAAGCGGTCGCGATC 910
Qy 241 CGCGGATCGATGAGCGCGCGCGCGACCGCGCGGCGGCGGCGGCGGCTGTCAGGTCGCG 300
Db 909 CGCGGATCGATGAGCGCGCGCGCGACCGCGCGGCGGCGGCGGCGGCTGTCAGGTCGCG 850
Qy 301 CGCGGATCGATGAGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 849 CGCGGATCGATGAGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 790
Qy 361 GAACAGCCCTGCG 420
Db 789 GAACAGCCCTGCG 730
Qy 421 TATGCGCTCATGAGGCGCGCGCGGATGAGGAGATGCGCGCTTTGCGCGCGCATCTCTGCGC 480
Db 729 TATGCGCTCATGAGGCGCGCGCGGATGAGGAGATGCGCGCTTTGCGCGCGCATCTCTGCGC 670
Qy 481 GGTCCAGCCCCGCTTGCGGCGATAGTCAGCGACGCTTGCGCGAGCGGCATCTGCGAG 540
Db 669 GGTCCAGCCCCGCTTGCGGCGATAGTCAGCGACGCTTGCGCGAGCGGCATCTGCGAG 610
Qy 541 ATCGCCGCGCTGCTGTAGCGCGTATCTCGATCAGATGCGGGTGGAGCTGAAGGGCAG 600
Db 609 ATCGCCGCGCTGCTGTAGCGCGTATCTCGATCAGATGCGGGTGGAGCTGAAGGGCAG 550
Qy 601 CAGATAGATGAAGCGGTACCGCTCCATCTGCGGAAAGCGTGCATCATATCATGCGGGG 660
Db 549 CAGATAGATGAAGCGGTACCGCTCCATCTGCGGAAAGCGTGCATCATATCATGCGGGG 490
Qy 661 CTCGACGCCATGAGGGGCGCTCGGCTCGATCTCGACGCGCCAGAAATTTCTGAAACCAC 720

Db 489 CTGACGCCATGAGGGGGGCTGCTGATCTGACGCCACGAAATTTGGAACCCAC 430
QY 721 GGTCAAGTGGGGGCTTCAGACGCAACCAAGGCGTGCATACGAGGACCTCGATCCG 780
Db 429 GGTCAAGTGGGGGCTTCAGACGCAACCAAGGCGTGCATACGAGGACCTCGATCCG 370
QY 781 CGAGCCGTCCGTAGCGTCCGCGCGGTATCTGCAGCGTGCAGCATGCGTATTCACCG 840
Db 369 CGAGCCGTCCGTAGCGTCCGCGCGGTATCTGCAGCGTGCAGCATGCGTATTCACCG 310
QY 841 CAATGACACCTCTGACGACGCCGATCAGCGCGCCCTCGATCGAGCATAGCTTGT 900
Db 309 CAATGACACCTCTGACGACGCCGATCAGCGCGCCCTCGATCGAGCATAGCTTGT 250
QY 901 CGTCAAGCGCGCGAATGCTCGGGAAACGCGACCTCTGATCCGTTCATTGCGCGAGCG 960
Db 249 CGTCAAGCGCGCGAATGCTCGGGAAACGCGACCTCTGATCCGTTCATTGCGCGAGCG 130
QY 961 AATGGCGACAGCGCGCGCATTCGGCGGAAAGTCCGTGTCGTGAGGACAGGT 1020
Db 189 AATGGCGACAGCGCGCGCATTCGGCGGAAAGTCCGTGTCGTGAGGACAGGT 130
QY 1021 GTCTGTGTCGAGGGGCGGACCGCGCGTGCAGCATCAGATGCGCGCATCCGCTTGGC 1080
Db 129 GTCTGTGTCGAGGGGCGGACCGCGCGTGCAGCATCAGATGCGCGCATCCGCTTGGC 70
QY 1081 GTGCGCAAGCGGCAAGCGCATCAGCGCACCGGACCGCGCGCGCATTCAGCGATC 1140
Db 69 GTGCGCAAGCGGCAAGCGCATCAGCGCACCGGACCGCGCGCGCATTCAGCGATC 10
QY 1141 ATGGCTCAT 1149
Db 9 ATGGCTCAT 1
RESULT 5
US-08-980-832-1
Sequence 1, Application US/08980832B
Patent No. 6291204
GENERAL INFORMATION:
APPLICANT: Paramontes, Luis
APPLICANT: Teygakov, Yuri
TITLE OF INVENTION: Improved Fermentative Carotenoid Production
FILE REFERENCE: Improved Fermentative Carotenoid
CURRENT APPLICATION NUMBER: US/08/980,832B
CURRENT FILING DATE: 1997-12-01
NUMBER OF SEQ ID NOS: 66
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 1
LENGTH: 8625
TYPE: DNA
ORGANISM: Flavobacterium sp. R1534
NAME/KEY: unsure
LOCATION: (8348) (8349)
NAME/KEY: unsure
LOCATION: (8539) (8540)
NAME/KEY: unsure
LOCATION: (8581)
NAME/KEY: unsure
LOCATION: (8590)
NAME/KEY: unsure
LOCATION: (8592)
NAME/KEY: unsure
LOCATION: (8602) (8604)
US-08-980-832-1

Query Match 100.0%; Score 1149; DB 3; Length 8625;
Best Local Similarity 100.0%; Pred. No. 1.3e-225;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGCTCTCTCTGAGAGAGGGGGCTTCGGGCAAGCGGACCGGCTTCGACAGCGG 60
Db 5794 TCATGCTCTCTCTGAGAGAGGGGGCTTCGGGCAAGCGGACCGGCTTCGACAGCGG 5853
QY 61 AATGGCGGGGCGTCCGCTGACATGCGAAGCCGGTGGGCAATGTGAGGGGCCGACATA 120
Db 5854 AATGGCGGGGCGTCCGCTGACATGCGAAGCCGGTGGGCAATGTGAGGGGCCGACATA 5913
QY 121 GAAAGCTCTGATCAGCGGCTTCGCGAGCGGTAAGAACCTCTGAGCAGCGATAGCGAGC 180
Db 5914 GAAAGCTCTGATCAGCGGCTTCGCGAGCGGTAAGAACCTCTGAGCAGCGATAGCGAGC 5973
QY 181 GTGAGGCGGGCAGCGCGGAAAGCATTCGGTTCAAGACCGGACCGGAGAAAGCGTCCGATC 240
Db 5974 GTGAGGCGGGCAGCGCGGAAAGCATTCGGTTCAAGACCGGACCGGAGAAAGCGTCCGATC 6033
QY 241 CGCGCATGATGATGCGCGCAGCGCGGACCGGCGAGCGGAGCGGAGCGGATCGATCGCG 300
Db 6034 CGCGCATGATGATGCGCGCAGCGCGGACCGGCGAGCGGAGCGGAGCGGATCGATCGCG 6093
QY 301 CGCGCATGATGATGCGCGCAGCGGATTCGCGGATAGGAGCAGATATCCGATGACGGGATG 360
Db 6094 CGCGCATGATGATGCGCGCAGCGGATTCGCGGATAGGAGCAGATATCCGATGACGGGATG 6153
QY 361 GAAAGCGCTGCGCCCGAGCGCAACCGGCAACCGGCGCGGCGGCTTCGCGGATCGGAGAGCC 420
Db 6154 GAAAGCGCTGCGCCCGAGCGCAACCGGCAACCGGCGCGGCGGCTTCGCGGATCGGAGAGCC 6213
QY 421 TATGCGCATGATGATGCGCGCAGCGGAGGAGGATGCGGCTTCGCGGCGGATCGGAGCC 480
Db 6214 TATGCGCATGATGATGCGCGCAGCGGAGGAGGATGCGGCTTCGCGGCGGATCGGAGCC 6273
QY 481 GGTCCAGCGCCCGCTGCGCGGATAGTCCAGGACAGCGCTTCGCGGACCGCATTCGTCAG 540
Db 6274 GGTCCAGCGCCCGCTGCGCGGATAGTCCAGGACAGCGCTTCGCGGACCGCATTCGTCAG 6333
QY 541 ATGCGCGCGCTGCTGTAGCGGATCTCTGATCAGATGCGGATGCGGATGCGGATGCGG 600
Db 6334 ATGCGCGCGCTGCTGTAGCGGATCTCTGATCAGATGCGGATGCGGATGCGGATGCGG 6393
QY 601 CAGATGATGAAAGCGGTACCGGTTCATTCGCGGAAAGCGTCCGATATCATTCGCGGCG 660
Db 6394 CAGATGATGAAAGCGGTACCGGTTCATTCGCGGAAAGCGTCCGATATCATTCGCGGCG 6453
QY 661 CTGACGCGCATGAGGGGCGTCCGATCTGATCTGACGCGCCACGAAATTTCTGAAACCCAC 720
Db 6454 CTGACGCGCATGAGGGGCGTCCGATCTGATCTGACGCGCCACGAAATTTCTGAAACCCAC 6513
QY 721 GGTCAAGTGGGGGCTTCAGACGCAACCAAGGCGTGCATCAGCGAGGACGCTTCGATCCG 780
Db 6514 GGTCAAGTGGGGGCTTCAGACGCAACCAAGGCGTGCATCAGCGAGGACGCTTCGATCCG 6573
QY 781 CGAGCCGTCCGTAGCGTCCGCGGTATCTGCAGCGTGCAGCATGCGTATTCACACCG 840
Db 6574 CGAGCCGTCCGTAGCGTCCGCGGTATCTGCAGCGTGCAGCATGCGTATTCACACCG 6633
QY 841 CAGATGACACCTCTGACGACGCCGATCAGCGCGCCCTCGATCGAGCATAGGCTTGT 900
Db 6634 CAGATGACACCTCTGACGACGCCGATCAGCGCGCCCTCGATCGAGCATAGGCTTGT 6693
QY 901 CGTCAAGCGCGCGAATGCTCGGGAAACGCGACTTCCTGATCCGTTCATTGCGCGAGCG 960
Db 6694 CGTCAAGCGCGCGAATGCTCGGGAAACGCGACTTCCTGATCCGTTCATTGCGCGAGCG 6753
QY 961 AATGGCGGACAGCGCGCGCAAGCATTCGGGCGAAAGATCCGTTGTCGAGAGAACAGGT 1020
Db 6754 AATGGCGGACAGCGCGCGCAAGCATTCGGGCGAAAGATCCGTTGTCGAGAGAACAGGT 6813
QY 1021 GTCTGTGTCGAGGGGCGGACCGCGCGTGCAGCATCAGATGCGCGCATCCGCTTGGC 1080

Accession	Sequence	Position
Db	6814 GTGCTGTGTCGAGGGGCGCCGGAACCGGCGGTTCAGAGATCAACGATGGGCGGATCCGGTCTTCG	6873
QY	1081 GTCCGGAAACGGCAACGGCGATCAGCGCACCCGACAGCCCGCGCCCGGATCAGACATC	1140
Db	6874 GTCCGGAAACGGCAACGGCGATCAGCGCACCCGACAGCCCGCGCCCGGATCAGACATC	6933
QY	1141 ATGGCTCAT 1149	
Db	6934 ATGGCTCAT 6942	

```

RESULT 6
US-09-923B-1
; Sequence 1, Application US/09920923B
; Patent No. 6677134
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; TITLE OF INVENTION: Fermentative Carotenoid Production
; FILE REFERENCE: 15464 US (C38435/125944)
; CURRENT APPLICATION NUMBER: US/09/920,923B
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

```

```

? ORGANISM: Flavobacterium sp. R1534
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8348)..(8349)
? OTHER INFORMATION: unsure
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8539)..(8540)
? OTHER INFORMATION: unsure
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8581)..(8581)
? OTHER INFORMATION: unsure
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8590)..(8590)
? OTHER INFORMATION: unsure
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8592)..(8592)
? OTHER INFORMATION: unsure
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (8602)..(8604)
? OTHER INFORMATION: unsure
? US-09-920-923B-1

```

Query Match	100.0%	Score 1149;	DB 4;	Length 8625;
Best Local Similarity	100.0%;	Pred. No. 1.3e-225;		
Matches 1149;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	Db	QY	Db
1	5794	61	5854
TCATGCTCTCTCTCAGACAGGGGGCGTTCCGGCAGGACGCGCCTTGGAACAGCGG	TCATGCTCTCTCTCAGACAGGGGGCGTTCCGGCAGGACGCGCCTTGGAACAGCGG	AATGGGGGGGGGCTCGGTGACGATGCGAAAGCCGGCTCGGCAATGTCAAGGCGCCGGACATTA	AATGGGGGGGGGCGCTCCGGTGACATGCGAAAGCCGGCTCGGCAATGTCAAGGCGCCGGACATTA
5853	5855	121	5914
GAAGCGCTTCGATCAGCGGCTGCGGCGAGGGCGTTAGAACCGCTTCACAGCAGGCGGATTAACGACG	GAAGCGCTTCGATCAGCGGCTGCGGCGAGGGCGTTAGAACCGCTTCACAGCAGGCGGATTAACGACG	GAAGCGCTTCGATCAGCGGCTGCGGCGAGGGCGTTAGAACCGCTTCACAGCAGGCGGATTAACGACG	GAAGCGCTTCGATCAGCGGCTGCGGCGAGGGCGTTAGAACCGCTTCACAGCAGGCGGATTAACGACG
180	5973	180	5973

QY	181	GTCCGGGGGGGAGCGCGGAAACAGCATCCGGTTCAGACGCGCAGAAAGCGTCCGATC	240
Db	5574	GTCCGGGGGGGAGCGCGGAAACAGCATCCGGTTCAGACGCGCAGAAAGCGTCCGATC	6033
QY	241	CGCGCGATCGATGGCGCCAGCGCGCACCGCGCAGCGGGCGGACGGGTCGTCAAGTCCGC	300
Db	6034	CGCGCGATCGATGGCGCCAGCGCGCACCGCGCGACCGGGCGGACGGGTCGTCAAGTCCGC	6093
QY	301	CGCGCGATGGCATTCGCGGACCTCCGGCGCATAGGGCAAGGGAATATCCGGTGAACGGGGTG	360
Db	6094	CGCGCGATGGCATTCGCGGACCTCCGGCGCGCATAGGGCAAGGGAATATCCGGTGAAGGGGTG	6153
QY	361	GAAACAGCCCTGGCCCCCAGGCCCAACCGGACCGGCCCCCTGGCGCGTGGTCCGCGCAGAAAGCC	420
Db	6154	GAAACAGCCCTGGCCCCCAGGCCCAACCGGACCGGCCCCCTGGCGCGTGGTCCGCGCAGAAAGCC	6213
QY	421	TATGGCGTCATGGGCGCAGCGCGATGGGCGAGGATGCCCTTTCGCGCGCCGACATTCCTGCC	480
Db	6214	TATGGCGTCATGGGCGCAGCGCGATGGGCGAGGATGCCCTTTCGCGCGCCGACATTCCTGCC	6273
QY	481	GGTCACAGCCCCGCTGGCGCGCATATGTCAGCGACGCTTCGGCCACGCGCGCCATTTGTCAG	540
Db	6274	GGTCACAGCCCCGCTGGCGCGCATATGTCAGCGACGCTTCGGCCACGCGCGCCATTTGTCAG	6333
QY	541	ATGCCCGCCCGTCCGTGAAGCGCGATCTCTGATCAAGATCGGGGTGGGACTGAAGGGGAG	600
Db	6334	ATGCCCGCCCGTCCGTGAAGCGCGATCTCTGATCAAGATGGGGGTGGGACTGAAGGGGAG	6393
QY	601	CAGATATGATGAAGCGGTACCCGTCATCTCGCGAAACGGTCCGCTCCATATCATCGGAGC	660
Db	6394	CAGATATGATGAAGCGGTATCCCGTCCATCTCGCGAAACGGTCCGCTCCATATCATCGGAGC	6453
QY	661	CTGCAGCCCATGGGGGGCGTCCGTCTCGATCTCGACGCGCCACGAAATTTCTGGAAACCCAC	720
Db	6454	CTGCAGCCCATGGGGGGCGTCCGTCTCGATCTCGACGCGCCACGAAATTTCTGGAAACCCAC	6513
QY	721	GCTCAAGGTGGGGGTCTCGACGGGACCACGGGCGTCGATCACGAGGACGCTCGATCCG	780
Db	6514	GCTCAAGGTGGGGGTCTCGACGGGACCACGGGCGTCGATCACGAGGACGCTCGATCCG	6573
QY	781	CGACCCGTCGTCAAGCGTCGCGCGGATCGTCAGCGTCGCGACGATGCGTATTCACCG	840
Db	6574	CGACCCGTCGTCAAGCGTCGCGCGGATCGTCAGCGTCGCGACGATGCGTATTCACCG	6633
QY	841	CAGATCGACACCTTCGACAGCAGCCCGGATCAAGCGCGCCGCTCGATCGAGCCATAGCTGT	900
Db	6634	CAGATCGACACCTTCGACAGCAGCCCGGATCAAGCGCGCCGCTCGATCGAGCCATAGCTGT	6693
QY	901	CGTCAGCGCGCGCGCAATATGTTGGGAAACGCGACTCTGATTCGTTCCATTCCGCGACG	960
Db	6694	CGTCAGCGCGCGCGCAATATGTTGGGAAACGCGACTCTGATTCGTTCCATTCCGCGACG	6753
QY	961	AATGGGCGACAGGGGCGCGCAGCCCATTTGGGGGAAAGATCCGTTGCTGGCGACAGAGT	1020
Db	6754	AATGGGCGACAGGGGCGCGCAGCCCATTTGGGGGAAAGATCCGTTGCTGGCGACAGAGT	6813
QY	1021	GTGCTGGTCCAGGGGCGCGGACCGCGCGTCGAGCATCAAGATGCGCGCATTCGGTCTGCG	1080
Db	6814	GTGCTGGTCCAGGGGCGCGGACCGCGCGTCGAGCATCAAGATGCGCGCATTCGGTCTGCG	6873
QY	1081	GTCCGGAACGGCAAGCGGATCAGCGCACCGGACAGCCCGCGCGCGATCAGCAGATC	1140
Db	6874	GTCCGGAACGGCAAGCGGATCAGCGCACCGGACAGCCCGCGCGCGATCAGCAGATC	6933
QY	1141	ATGGCTCAT 1149	
Db	6934	ATGGCTCAT 6942	

RESULT 7
US-08-980-832-27
; Sequence 27, Application US/089808032B
; Patent No. 6291204

```

; GENERAL INFORMATION:
; APPLICANT: Tsigankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/08/980,832B
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 11233
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Plasmid pZee4
US-08-980-832-27

Query Match      100.0%; Score 1149; DB 3; Length 11233;
Best Local Similarity 100.0%; Pred. No. 1.4e-225;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGCTCTCTCTGAGAGAGGGGGCGTTTCGGGACAGCGACGCGCTTCGACAGCGG 60
DB 6476 TCATGCTCTCTCTGAGAGAGGGGGCGTTTCGGGACAGCGACGCGCTTCGACAGCGG 6535
QY 61 AATGGGCGGGCGCTCGGTGACGATGCGAAGCCGGTGGCGCAATGTCAGGCGCCCGGCATTA 120
DB 6536 AATGGGCGGGCGCTCGGTGACGATGCGAAGCCGGTGGCGCAATGTCAGGCGCCCGGCATTA 6595
QY 121 GAAGCGCTGATGAGCGGCTGCGGACAGCGGTAGAAACCGCTGACAGAGCGATAGCGACG 180
DB 6596 GAAGCGCTGATGAGCGGCTGCGGACAGCGGTAGAAACCGCTGACAGAGCGATAGCGACG 6655
QY 181 GTGCGGCGGCGACGCGCGGAAACGATCCGGTTTCAGACGCGGAGAAAGCGGTGCGGATC 240
DB 6656 GTGCGGCGGCGACGCGCGGAAACGATCCGGTTTCAGACGCGGAGAAAGCGGTGCGGATC 6715
QY 241 CGCGCGATGATGAGCGGCGGACGCGCGGACGCGCGGAGAGCGGCGGTGTCAGGTCGCG 300
DB 6716 CGCGCGATGATGAGCGGCGGACGCGCGGACGCGCGGAGAGCGGCGGTGTCAGGTCGCG 6775
QY 301 CGCGCGATGATGAGCGGCGGACGCGCGGATGAGGAGCGGATATCCGGTGAAGCGGATG 360
DB 6776 CGCGCGATGATGAGCGGCGGACGCGCGGATGAGGAGCGGATATCCGGTGAAGCGGATG 6835
QY 361 GAACAGCGCTGCGCGCGGACGCGGACGCGCGGCGGCGGCGGTGTCGCGCGGAGAGCG 420
DB 6836 GAACAGCGCTGCGCGCGGACGCGGACGCGCGGCGGCGGCGGTGTCGCGCGGAGAGCG 6895
QY 421 TATGGGCTGATGAGCGGCGGACGCGGATGAGGAGCGGATGCGCGGCGGATCTCTGCGCC 480
DB 6896 TATGGGCTGATGAGCGGCGGACGCGGATGAGGAGCGGATGCGCGGCGGATCTCTGCGCC 6955
QY 481 GGTTCAGCGCGCGCTGCGGCGGATGATGACGCGCGCTGCGGCGGCGGCGGCGGATGTCAG 540
DB 6956 GGTTCAGCGCGCGCTGCGGCGGATGATGACGCGCGCTGCGGCGGCGGCGGCGGATGTCAG 7015
QY 541 ATCGCGCGCGCTGCGGCGGATGATGACGCGCGCTGCGGCGGCGGCGGCGGCGGATGTCAG 600
DB 7016 ATCGCGCGCGCTGCGGCGGATGATGACGCGCGCTGCGGCGGCGGCGGCGGCGGATGTCAG 7075
QY 601 CAGATGATGAGAGCGGATGCGCGGATGATGCGGAGAGCGGCGGCGGATGATGTCAGGCGG 660
DB 7076 CAGATGATGAGAGCGGATGCGCGGATGATGCGGAGAGCGGCGGCGGATGATGTCAGGCGG 7135
QY 661 CTCGAGCGCGGCGGCGGCGGATGATGCGGAGAGCGGCGGCGGCGGATGATGTCAGGCGG 720
DB 7136 CTCGAGCGCGGCGGCGGCGGATGATGCGGAGAGCGGCGGCGGCGGATGATGTCAGGCGG 7195
QY 721 GGTTCAGGTCGCGGCGGTCGAGCGGACGAGCGGCGGCGGATGATGTCAGGCGGAGCGCTGATCG 780
DB 7196 GGTTCAGGTCGCGGCGGTCGAGCGGACGAGCGGCGGCGGATGATGTCAGGCGGAGCGCTGATCG 7255

```

```

QY 781 CGAGCGCTGCTGAGCGGTCGCGCGGATGCTGACAGCGTGGAGATGCGATTCACCG 840
DB 7256 CGAGCGCTGCTGAGCGGTCGCGCGGATGCTGACAGCGTGGAGATGCGATTCACCG 7315
QY 841 CAGATGACACCTGTCAGAGAGCGGTCGAGCGGCGGCGGCGGTCGATGAGCGGATGCTG 900
DB 7316 CAGATGACACCTGTCAGAGAGCGGTCGAGCGGCGGCGGCGGTCGATGAGCGGATGCTG 7375
QY 901 CGTCAGCGGCGGCGGATGCTGCGGAGAAAGCGGACCTTCGATGCGGCGGCGGAG 960
DB 7376 CGTCAGCGGCGGCGGATGCTGCGGAGAAAGCGGACCTTCGATGCGGCGGCGGAG 7435
QY 961 AATGGGCGGACAGCGGCGGCGGCGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 7436 AATGGGCGGACAGCGGCGGCGGCGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7495
QY 1021 GTGCTGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 7496 GTGCTGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7555
QY 1081 GTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 7556 GTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7615
QY 1141 ATGGGCTCAT 1149
DB 7616 ATGGGCTCAT 7624

RESULT 8
US-09-920-923B-27
; Sequence 27, Application US/09920923B
; Patent No. 6677134
; GENERAL INFORMATION:
; APPLICANT: Tsigankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/09/920,923B
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 11233
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Plasmid pZee4
US-09-920-923B-27

Query Match      100.0%; Score 1149; DB 4; Length 11233;
Best Local Similarity 100.0%; Pred. No. 1.4e-225;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGCTCTCTCTGAGAGAGGGGGCGTTTCGGGACAGCGACGCGCTTCGACAGCGG 60
DB 6476 TCATGCTCTCTCTGAGAGAGGGGGCGTTTCGGGACAGCGACGCGCTTCGACAGCGG 6535
QY 61 AATGGGCGGGCGCTCGGTGACGATGCGAAGCCGGTGGCGCAATGTCAGGCGCCCGGCATTA 120
DB 6536 AATGGGCGGGCGCTCGGTGACGATGCGAAGCCGGTGGCGCAATGTCAGGCGCCCGGCATTA 6595
QY 121 GAAGCGCTGATGAGCGGCTGCGGACAGCGGTAGAAACCGCTGACAGAGCGGATAGCGACG 180
DB 6596 GAAGCGCTGATGAGCGGCTGCGGACAGCGGTAGAAACCGCTGACAGAGCGGATAGCGACG 6655
QY 181 GTGCGGCGGCGACGCGCGGAAACGATCCGGTTTCAGACGCGGCGGCGGAGAGCGGTGCGGATC 240
DB 6656 GTGCGGCGGCGACGCGCGGAAACGATCCGGTTTCAGACGCGGCGGCGGAGAGAGCGGTGCGGATC 6715
QY 241 CGCGGATGATGAGCGGCGGACGCGCGGAGAGCGGCGGAGAGCGGCGGCGGATGTCAGGTCGCG 300

```

Db	6716	CGCGGATTCGATGGCGCCAGCCCGCGGACCCGCGGACCGCGCGGACCGCGGTCGTCAGGTCCG	6775
Qy	301	CGCGCGATGGGATTCGCGGACCTTGCGCGGCATTAAGGGCAGCGAATATCCGCTGACGGGGTG	360
Db	6776	CGCGCGATGGGATTCGCGGACCTTGCGCGGCATTAAGGGCAGCGAATATCCGCTGACGGGGTG	6835
Qy	361	GAAACGCCCTGCCCCCAGCGCCAAACCGGACCGCGCCCCCGGGCGGTGTGGGCGCAGAAAGCC	420
Db	6836	GAAACGCCCTGCCCCCAGCGCCAAACCGGACCGCGCCCCCGGGCGGTGTGGGCGCAGAAAGCC	6895
Qy	421	TATGGCGTCATAGGAGCAGCGGATGGGAGGATGCCCTTTTGCGCGCGCATCTCTTGCC	480
Db	6896	TATGGCGTCATAGGAGCAGCGGATGGGAGGATGCCCTTTTGCGCGCGCATCTCTTGCC	6955
Qy	481	GATCCAGCCCCGCTTGCGCGCATATGTCAGCGAGCGCTTGCGCCAGCGCGCATGCTCCAG	540
Db	6956	GATCCAGCCCCGCTTGCGCGCATATGTCAGCGAGCGCTTGCGCCAGCGCGCATGCTCCAG	7015
Qy	541	ATCGCGCGCGCTGCGGTGAGCGCGATATCTTCGATCAGAGATGGGGGTGGGACTGAAGGGGAG	600
Db	7016	ATCGCGCGCGCTGCGGTGAGCGCGATATCTTCGATCAGAGATGGGGGTGGGACTGAAGGGGAG	7075
Qy	601	CAGATAGATGAAGCGGTACCGGTCATCTGCGGAAACGCTGCGCTCATGATCATTCGAGCG	660
Db	7076	CAGATAGATGAAGCGGTACCGGTCATCTGCGGAAACGCTGCGCTCATGATCATTCGAGCG	7135
Qy	661	CTCGACGCCATGGGGGGGCGTGGGTCTCGATCTCGACGCCCAAGAAATTTCTGAAACCAC	720
Db	7136	CTCGACGCCATGGGGGGGCGTGGGTCTCGATCTCGACGCCCAAGAAATTTCTGAAACCAC	7195
Qy	721	GCTCAGGTGCGGGGCTCTCGACGGCACCAACCGGCGTTCATTCACGACGAGGAGCTTCGATCCG	780
Db	7196	GCTCAGGTGCGGGGCTCTCGACGGCACCAACCGGCGTTCATTCACGACGAGGAGCTTCGATCCG	7255
Qy	781	CGAGCCGCTCCGTCAGCGCTGCGCGCGGTATGTCAGCGCTGCGGACATGCTATTCCACCG	840
Db	7256	CGAGCCGCTCCGTCAGCGCTGCGCGCGGTATGTCAGCGCTGCGGACATGCTATTCCACCG	7315
Qy	841	CAGATCGACACCTTGCGACGAGCCCGATCAGCGCGCGCGCTTCGATCGAGCCATAGCTGT	900
Db	7316	CAGATCGACACCTTGCGACGAGCCCGATCAGCGCGCGCGCTTCGATCGAGCCATAGCTGT	7375
Qy	901	CGTCAAGCGCGCGCGGAATGTGTGGGAAACCGGACCTTCGATTCGCTTCATTGGCGCGACG	960
Db	7376	CGTCAAGCGCGCGCGGAATGTGTGGGAAACCGGACCTTCGATTCGCTTCATTGGCGCGACG	7435
Qy	961	AATGGGCGACAGGCGCGCGCAGCCATTTGCGGGCGAAAGATCCCGGTGTGTGGGCGAGCACAGT	1020
Db	7436	AATGGGCGACAGGCGCGCGCAGCCATTTGCGGGCGAAAGATCCCGGTGTGTGGGCGAGCACAGT	7495
Qy	1021	GTGCTGTGTCGAGGGGCGCGAACCGGCGCTCGAGCATCAAGATGGCGCGATCCGCTTGCG	1080
Db	7496	GTGCTGTGTCGAGGGGCGCGAACCGGCGCTCGAGCATCAAGATGGCGCGATCCGCTTGCG	7555
Qy	1081	GTCGCGAAGCGCAGCGCGATCAGCGCACCGGACAGCCCCCGCGCGCGGATCAGCATC	1140
Db	7556	GTCGCGAAGCGCAGCGCGATCAGCGCACCGGACAGCCCCCGCGCGCGGATCAGCATC	7615
Qy	1141	ATGGCTCAT 1149	
Db	7616	ATGGCTCAT 7624	

```

1  TITLE OF INVENTION:  DNA STRADS USEFUL FOR THE SYNTHESIS OF
2  TITLE OF INVENTION:  XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
3  TITLE OF INVENTION:  XANTHOPHYLLS
4  NUMBER OF SEQUENCES:  12
5  CORRESPONDENCE ADDRESS:
6  ADDRESSER:  Foley & Lardner
7  STREET:  3000 K Street, N.W., Suite 500
8  CITY:  Washington
9  STATE:  D.C.
10 COUNTRY:  USA
11 ZIP:  20007-5109
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  PatentIn Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/08/663,310
19 FILING DATE:  23-SEP-1996
20 CLASSIFICATION:  435
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER:  WO PCT/JP94/02220
23 FILING DATE:  26-DEC-1994
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER:  JP 6-235917
26 FILING DATE:  05-SEP-1994
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER:  JP 5-348737
29 FILING DATE:  27-DEC-1993
30 ATTORNEY/AGENT INFORMATION:
31 NAME:  BENT, Stephen A.
32 REGISTRATION NUMBER:  29,768
33 REFERENCE/DOCKET NUMBER:  49441/109
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE:  (202)672-5300
36 TELEFAX:  (202)672-5399
37 TELEX:  904136
38 INFORMATION FOR SEQ ID NO:  5:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH:  1161 base pairs
41 TYPE:  nucleic acid
42 STRANDEDNESS:  double
43 TOPOLOGY:  linear
44 FEATURE:
45 NAME/KEY:  CDS
46 LOCATION:  1..1158
47 FEATURE:
48 NAME/KEY:  mat peptide
49 LOCATION:  1..1158
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922

```

Db	921	GGCCCGGTGANTGGGTATATGGCGGATGAGCGCGCGCAGCGCGTGTGGTCCGGCGCGCC	862
Oy	235	GTCCGCGCGCCGCGATGCGATCCGCGACTCTGCGCGGCATATGGCGACGGAATATCCGTGTAC	354
Db	861	GGACAGACCCCGCACCAAGTCGCGCACCTGTGTCCCATATAGGGCAGCGAATATGCCGTGTAC	802
Oy	355	GGGGGTGGAACAGCCCCGTGCCCCAGCGCCAAACCGGGACCGCCCCCTTGGGGTGGTCCGGCA	414
Db	801	CGGATGAAAGAACCCCGCGCGGATGCCACCGGAAACAGGCCCCCGCGGTGATCCGCCCA	742
Oy	415	GAAGCCTATGCGGTCTATGGGCGAGCGCGCATGGGCGAGATGCCCCCTTTGCGCGCGCATCTC	474
Db	741	GAAGCCCGCGCGATCATGGGCGCAGCGCGATGGGAGAGATGCCGCGTTGCGCGCGGACCTTC	682
Oy	475	CTGCGCGGTTCAGCCCCCGCTTGGCGCGCATATGTCACGCGACGCTTGGCGCAGCGCGCATTC	534
Db	681	GGCCCCGGTCCAGCCCTCTGTGGCGGGCATATGTCGTGGAGACGCGCGCGCAGGCGCTCGTC	622
Oy	535	GTCCAGATTCGCGCGCGGTGCGTGTACGGGTATCTCGATTCAGGATAGGGGTGGGACTGAA	594
Db	621	GTCCAGATTCGCGCGCGCATGAGAAATACGGGTGTCTCGATTCAGGAATGCGCGTGGAGAGAA	562
Oy	595	GGGCGACGATAGATGAAGACGGGTACCCGTCCATCTTGCAGAAACGCTGCGGTCCATGATCAT	654
Db	561	GGGCGACGATAGATGAAGACGGGTACCCGTCTCTGCGGTGAGACGCTGCGGTCCATGATCAT	502
Oy	655	CGGGCGCTCGACGCCATGAGGGGGCGTGGGTCTCGATCTCGACGCCCGACGAATTTCTGAAA	714
Db	501	CGGGCGGGGCAAGCGGTGAGGGCGGTGCGGTCTCGATCTCGACGCCCGACGAATTTCTGAAA	442
Oy	715	ACCCACGGTCAAGTGGGGGGTTCGACCGGCATCACGGGCGTTCGATCAACGAGCGACGCTTC	774
Db	441	ACCCACGGTCAAGTGGGGGGTTCGACCGGTGCGCGCCCGCGCTTCAGAGACCGCGCGCTTC	382
Oy	775	GATCCGCGAGCCGTCCTGTACGCGTGCAGCGCGGTATCTGTACGCGTCCGAGCATGCGTAT	834
Db	381	GATCCGCGGTGCGCGAGACAGCGTGCCTCCCTCTGCGCATTCAGACAGGCGAGATGTGCTGTCC	322
Oy	835	CCACCGCAGATCGACACCTGTGACGACG-----CGATTCAGCGCGCGCCGCTCGATCGA	888
Db	321	CCAGCGGATCTTCGGCGCCGACCGGACCAACCGCATTCGCGCAGCGCGCGCCGCTTCAGGGA	262
Oy	889	GCATATGCTCTGTCTGCAGGCGCGCGGAATGCTCGGGAACCGGACCTTCTGATCCGTCTCA	948
Db	261	CCCGTAAACGGGTGGCGCACCGCGCGGCGATGGCGGGGAAGCGCACCTCTGGTCCGGGCA	202
Oy	949	TTCCGCGGAGGAATGGGCGACAGGCGCGGCGACGCAATTCGGGCGAAATCCGTGTCTGTG	1008
Db	201	GTGGCGGGCGGACAGGGGCTTTCAGCGCGGCGACGATTCGCGCGACAGGTTCGGGTGTGTG	142
Oy	1009	GCAGGACCGAGGTGTGTCTGTGCTCGAGAGGGGCGGACCGCGCTGTGACGATTCAGATCGCGC	1068
Db	141	GCAAGGACCAAGGTGTGGCGGTCTGAAGGTCTGTGGCGCATGTCTCAGACACGACGACGCGAG	82
Oy	1069	ATCCGGTCTGGGTGCGTCCGACAGCGCGCATTCAGCGCACCGGACAGCGCCCGCGCCGCG	1128
Db	81	GTCCGGCGCGCGCGCGCGCGCAGCGCGCATTCAGCCGTTGGCGAAGCGCGCGCCCTTGC	22
Oy	1129	GATCAGCAGATCATGCTCTCA 1148	
Db	21	CAGCAGCACGTCAATGGGTCA 2	

RESULT 10
 US-09-006-491-5/c
 ; Sequence 5, Application US/09006491
 ; Patent No. 5972690
 ; GENERAL INFORMATION:
 ; APPLICANT: MISAWA, No. 5972690jhiiko
 ; APPLICANT: KONDO, Keiji
 ; APPLICANT: KAJIWARA, Susumu
 ; APPLICANT: YOKOYAMA, Akihito
 ; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF

```

1  TITLE OF INVENTION:  XANTHOPHILS AND THE PROCESS FOR PRODUCING THE
2  TITLE OF INVENTION:  XANTHOPHILLS
3  NUMBER OF SEQUENCES:  12
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE:  Foley & Lardner
6  STREET:  3000 K Street, N.W., Suite 500
7  CITY:  Washington
8  STATE:  D C
9  COUNTRY:  USA
10 ZIP:  20007-5109
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  Patentln Release #1.0, Version #1.30
17
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER:  US/09/006,491
20 FILING DATE:
21
22 CLASSIFICATION:
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER:  US 08/663,310
26 FILING DATE:  23-SEP-1996
27 APPLICATION NUMBER:  WO PCT/JP94/02220
28 FILING DATE:  26-DEC-1994
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER:  JP 6-235917
31 FILING DATE:  05-SEP-1994
32
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER:  JP 5-348737
35 FILING DATE:  27-DEC-1993
36 ATTORNEY/AGENT INFORMATION:
37
38 NAME:  BENT, Stephen A.
39 REGISTRATION NUMBER:  29,768
40 REFERENCE/DOCKET NUMBER:  49441/109
41 TELECOMMUNICATION INFORMATION:
42 TELEPHONE:  (202) 672-5300
43 TELEFAX:  (202) 672-5399
44
45 TELEX:  904136
46 INFORMATION FOR SEQ ID NO:  5:
47
48 SEQUENCE CHARACTERISTICS:
49 LENGTH:  1161 base pairs
50 TYPE:  nucleic acid
51 STRANDEDNESS:  double
52 TOPOLOGY:  linear
53
54 FEATURE:
55
56 NAME/KEY:  CDS
57 LOCATION:  1..1158
58 FEATURE:
59
60 NAME/KEY:  mat_peptide
61 LOCATION:  1..1158
62
63 US-09-006-491-5
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
91
```

DB 921 CCCCCGATGATGCGTATCGCGATGCGCGCGAGCGCGTGCCTCCGGCGCGCC 862
QY 295 GTCCGCGCGCGCGATGCGATCCGCGACCTGCGCGGATGAGGACAGAAATCCGCTAC 354
DB 861 GACAGACCCCGCACCACTCCGCCACTGTCCGATGAGGACAGAAATACCGGTGAC 802
QY 355 GGGGTGAAACAGCCCTGCCCCAGCCGCAACCGGACCGCCCTGCGCGTGTCCGCCA 414
DB 801 CGGATGAAAGAACCCCGCGCGCATGCCACGGGAACAGGCCCGCGCGCTGATCCGCCA 742
QY 415 GAAGCTTATGCGCTATGCGGCAAGCGCGATGCGGAGATGCCCTTTCCGCGCATCTC 474
DB 741 GAAGCCCGCGCATGATGCGGCAAGCGCGATGCGGAGATGCCCGCTTCCGCGCATCTC 682
QY 475 CTGCG 534
DB 681 GCG 622
QY 535 GTCCAGATGCG 594
DB 621 GTCCAGATGCG 562
QY 595 GGGCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 654
DB 561 GGGCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
QY 655 GCG 714
DB 501 CCG 442
QY 715 ACCCAGCGTCAAGTGTGCGGCTGTCCAGCGCGCACCGCGCGCGCGCGCGCGCGCTC 774
DB 441 ACCCAGCGTCAAGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 382
QY 775 GATCCG 834
DB 381 GATCCG 322
QY 835 CCAGCGCGATGACACCGCTGACGAGC-----CCGATCAGCGCGCGCGCGCGCGCG 888
DB 321 CCAGCGCGATGACACCG 262
QY 889 GCCATGCGCTGTCTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948
DB 261 CCGGTACCGGTGTGCG 202
QY 949 TTGCG 1008
DB 201 GTTGGCG 142
QY 1009 GAGAGCG 1068
DB 141 GAGAGCG 82
QY 1069 ATCCGCTGTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1128
DB 81 GTCCG 22
QY 1129 GATCAGCGATCATGCGCTCA 1148
DB 21 CAGCAGCAGCATGATGATGATCA 2

RESULT 11

US-09-335-919-5/c
; Sequence 5, Application US/09335919
; Patent No. 6150130
; GENERAL INFORMATION:
; APPLICANT: MISAWA, No. 6150130.hiko
; APPLICANT: KONDO, Keiji
; APPLICANT: KAJIWARA, Susumu
; APPLICANT: YOKOYAMA, Akihito

;/ TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
;/ TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
;/ TITLE OF INVENTION: XANTHOPHYLLS
;/ NUMBER OF SEQUENCES: 12
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSES: Foley & Lardner
;/ STREET: 3000 K Street, N.W., Suite 500
;/ CITY: Washington
;/ STATE: D.C.
;/ COUNTRY: USA
;/ ZIP: 20007-5109
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ OPERATING SYSTEM: IBM PC compatible
;/ SOFTWARE: Patent Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/335,919
;/ FILING DATE:
;/ CLASSIFICATION:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/663,310
;/ FILING DATE: 23-SEP-1996
;/ APPLICATION NUMBER: WO PCT/JP94/02220
;/ FILING DATE: 26-DEC-1994
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: JP 6-235917
;/ FILING DATE: 05-SEP-1994
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: JP 5-348737
;/ FILING DATE: 27-DEC-1993
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: BENT, Stephen A.
;/ REGISTRATION NUMBER: 29,768
;/ REFERENCE/DOCKET NUMBER: 49441/109
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (202)672-5300
;/ TELEFAX: (202)672-5399
;/ TELEX: 904136
;/ INFORMATION FOR SEQ ID NO: 5:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1161 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
;/ NAME/KEY: CDS
;/ LOCATION: 1..1158
;/ FEATURE:
;/ NAME/KEY: mat_peptide
;/ LOCATION: 1..1158
;/ US-09-335-919-5
Query Match 55.8%; Score 640.8; DB 3; Length 1161;
Best Local Similarity 73.4%; Pred. No. 5.6e-122;
Matches 851; Conservative 0; Mismatches 297; Indels 12; Gaps 2;
QY 1 TCATGCTCTCTCTCTCAG 60
DB 1161 TCATGCTCTCTCTCTCAG 1102
QY 61 AATGGCGCGCGCGTCCGATGACGATGCGAAGCGCGTCCGATGTCAGGCGCGCGCATTA 120
DB 1101 AATGGAGAGCGTTGCCGTCACGATGCGAGCTGATCCGCCACGCTCACCGCGCGCATTA 1042
QY 121 GAAGCGCTGATCAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 1041 GAAGCGCTGATCAGCTCCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 982
QY 181 GTCCG 240
DB 981 GTCCG 922

```

QY 241 CCGCGATCGATGAGCCCGGCGACCGCGCAGC-----GGCGAGCGCGTCTGAC 294
DB 921 CCGCCGATGATCGCGTAATCGGAGATGGCGCGCGCGGCGCGTCCGTGCGCGCGGCC 862
QY 295 GTGCGCGCGCGCGATGCGATCCCGGACCTGCGCGGATAGGAGAGGAATATCCGCTAC 354
DB 861 GAGCAGACCCCGCCACGATCCCGCACCTGTGCGGATAGGAGAGGAATAGCGGTGAC 802
QY 355 GGGGTAACAGAGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 414
DB 801 CGGATAGAAAGAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 742
QY 415 GAAAGCTATGCGCTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
DB 741 GAAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 682
QY 475 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
DB 681 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 622
QY 535 GTCCAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
DB 621 GTCCAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 562
QY 595 GGGCAGCAGATAGATAGAGCGGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654
DB 561 GGGCAGCAGATAGATAGAGCGGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 502
QY 655 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
DB 501 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 442
QY 715 ACCCAGCGTCAAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 774
DB 441 ACCCAGCGTCAAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 382
QY 775 GATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 834
DB 381 GATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 322
QY 835 CCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 888
DB 321 CCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 262
QY 889 GCCATAGCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948
DB 261 CCGTAACCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 202
QY 949 TTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1008
DB 201 GTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 142
QY 1009 GAGAGCAGGTGTGCTGCTCGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1068
DB 141 GAGAGCAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 82
QY 1069 ATCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1128
DB 81 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22
QY 1129 GATCAGCAGATCATGCTCA 1148
DB 21 CAGCAGCAGCTCATGCTCA 2

```

```

APPLICANT: YOKOYAMA, Akihito
TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
TITLE OF INVENTION: XANTHOPHYLLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,310
FILING DATE: 23-SEP-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/02220
FILING DATE: 26-DEC-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-235917
FILING DATE: 05-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 5-348737
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2886 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-663-310-7
Query Match 55.8%; Score 640.8; DB 1; Length 2886;
Best Local Similarity 73.4%; Pred. No. 6,1e-122;
Matches 851; Conservative 0; Mismatches 297; Indels 12; Gaps 2;
QY 1 TCATGCTCTCTCTGACGAGGCGCGTTCGGGCGAGCGCGCGCGCGCGCGCGCGCG 60
DB 2509 TCATGCTCTCTCTCTGACGAGGCGCGTTCGGGCGAGCGCGCGCGCGCGCGCG 2450
QY 61 AATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 2449 AATGGAGGCGTTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2390
QY 121 GAAGCGCTGATCAGCGCGCTGCGGCGAGCGGTGAGAACCGCTGCGAGCGAGCGAG 180
DB 2389 GAACGCTTGCATGATGTCATGCGGCGATGCGGTAGAACCGCTGCGAGCGAGCG 2330
QY 181 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 2329 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2270
QY 241 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 294
DB 2269 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2210
QY 295 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 354
DB 2209 GAGCAGACCCCGCCACGATCCCGCACCTGTGCGGATAGGAGAGGAATAGCGGTGAC 2150

```

OY	355	GGGGTGAAGAACG	CCCTGGCCCCCAGCCGCAACCGGACACGCCCCCTTGCGGTGATCGCGCA	414
Db	2149	CGGATGAAGAAAC	CCCCCGCCGACGATGCCACCGGAAACAGCCCCCGCGCGTATCGGACCA	2099
OY	415	GAAAGCTTATGAG	CGTCATGAGGCCAGCGACGATAGGACAGATGCCCCCTTTCGCGCCGATCTC	474
Db	2089	GAAAGCCCGCCG	ATCATGAGGACGACGCGATGAGAGATGACCGCGCTTTCGCGCCGACCTTC	2030
OY	475	CTGCCCCGCTCA	GCCCCCTTGGCCCCGCGCATATGTCACAGCAACGCTTCGCGCAAGCGGCATTC	534
Db	2029	GGCCCCGGGTCA	CGCCCTCTCTGGCCGGGCATATGTCGAGGACCGCGCGCCAGCGCTGTCTC	1970
OY	535	GTCCAGATTCG	CCGCGCGCTGTGACCTGTACCGGATATCCTCGATACAGATATCGAGACTGAA	594
Db	1969	GTCCAGATTCG	CCGCGCGCATGCGGATATGACGCGGTCTCTCGATACAGATATGCCGCTCGAGAGA	1910
OY	595	GGGACGACAGAT	ATGATGAACCGGTACCCGCTCATCTTCGAGAACGAGTCGCGTCCATGATCAT	654
Db	1909	GGGACGACAGAT	ATGATGAACCGGTATACCCGCTCATGCTGGGTGACGCGTCCGCTCATGATCAT	1850
OY	655	CGGCGGCTCGA	ACGCCATAGGGGGGCGCTCGGTCTTCGATCTTCGACGCCCAAGAAATTTGTGAA	714
Db	1849	CGGCGGCGAC	ACGCCGTGCGGCGCGGTCTCGATCTTCGATCTTCGACCCCAAGAAATTTGTGAA	1790
OY	715	ACCCACGCTCA	GATCGGAGTCTCGAGGAGTCTCGACAGGACACACAGGCGGTGATCAACGACAGCA	774
Db	1789	ACCCACGCTCA	GATCGGAGTCTCGGCGATCCCGCGACGCGTGGCGCGCCCGCGCTTCAGAGAC	1730
OY	775	GATCCGCGA	CGCCGTCCGTGACGCTCGCGCGGATGTCACGCGTCGCGACATGCGTATT	834
Db	1729	GATCCGCGGTG	CCGCGACAGACGCGTCCGCCCCCTTCGCGCATTCAGACAGGCGAGATGCTGTCTC	1670
OY	835	CCACCGCGAT	TGACACCTTCGACGAGC-----CCGATTCAGCGCGCCCGCGCTCGATGGA	888
Db	1669	CCAGCGGAT	CTTGCGCGCCCGACCGGACCAACGCGATTCGCGACGCGCGCCGCTTCAGCGGA	1610
OY	889	GCCATAGCCT	TGCTCTGACGCGCGCGCAATGATCGGAGAAACGCGACCTCTGATCCGTCA	948
Db	1609	CCCGTAA	CCGGTGGCGACGCGCGCGGCAATGCGGGGAGAAACGCGACCTCTGGTGGGCGCA	1550
OY	949	TTTCGCGCGA	TCGATAGGCGCGACAGCGCGCGCCAGCGCATTCGGCGGAAAGATCCGTCTCTG	1006
Db	1549	GTTTCGCGCG	CGCCACAGGCGGCTTCAGCCGCGCCAGCCAGTCCCGCGCACAGGTCCGGGTCGTG	1490
OY	1009	GCAGGACCA	GATGATCTGTCTCCGAGGAGGCGCGACCGCGGTGAGACATCAAGATCGCGCG	1066
Db	1489	GAGGACCA	GATGATCTGTCTTCGAGGATCTTCGCGCATGATTCAGGACGACAGCGCGAG	1430
OY	1069	ATCCGCTG	TCGATGCGCGAACGCGACGCGATTCAGCGCACCGGACAGCCCCGCGCGCG	1126
Db	1429	GTCCGCGCG	CGCGCGCGCGACGCGCGCATGACCGCGTTGGAGAAAGGCCCGCGCTGCG	1370
OY	1129	GATCAGCAGAT	CATGAGCTCA 1148	
Db	1369	CAGCAGCA	CGTCACTGAGTCA 1350	
RESULT 13				
US-09-006-491-7/c				
; Sequence 7, Application US/09006491				
; Patent No. 5972690				
; GENERAL INFORMATION:				
; APPLICANT: MISAWA, No. 5972690ihiko				
; APPLICANT: KONDO, Keiji				
; APPLICANT: KAJIWARA, Susumu				
; APPLICANT: YOKOYAMA, Akihiro				
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF				
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE				
; NUMBER OF SEQUENCES: 12				
; CORRESPONDENCE ADDRESS:				
; ADDRESS: Foley & Lardner				
;				

```

1 STREET: 3000 K Street, N.W., Suite 500
2 City: Washington
3
4 STATE: D.C.
5
6 COUNTRY: USA
7 ZIP: 20007-5109
8
9 COMPUTER READABLE FORM:
10 MEDIUM TYPE: Floppy disk
11
12 COMPUTER: IBM PC compatible
13 OPERATING SYSTEM: PC-DOS/MS-DOS
14 SOFTWARE: Patent In Release #1.0, Version #1.30
15
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/09/006,491
18 FILING DATE:
19
20 CLASSIFICATION:
21
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 08/663,310
24 FILING DATE: 23-SEP-1996
25
26 APPLICATION NUMBER: WO PCT/JP94/02220
27 FILING DATE: 26-DEC-1994
28
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: JP 6-235917
31 FILING DATE: 05-SEP-1994
32
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: JP 5-348737
35 FILING DATE: 27-DEC-1993
36
37 ATTORNEY/AGENT INFORMATION:
38 NAME: BENT, Stephen A.
39
40 REGISTRATION NUMBER: 29,768
41
42 REFERENCE/DOCKET NUMBER: 49441/109
43
44 TELECOMMUNICATION INFORMATION:
45 TELEPHONE: (202) 672-5300
46
47 TELEFAX: (202) 672-5399
48
49 TELEX: 904136
50
51 INFORMATION FOR SEQ ID NO. 7:
52
53 SEQUENCE CHARACTERISTICS:
54 LENGTH: 2886 base pairs
55 TYPE: nucleic acid
56 STRANDEDNESS: double
57
58 TOPOLOGY: linear
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957

```

Query Match	55.8%	Score 640.8	DB 2	Length 2886
Best Local Similarity	73.4%	Pred. No. 6.1e-122		
Matches	851	Conservative	0	Mismatches 297
			Indels	12
			Gaps	2
QY	1	TCATGCTTCTCTCCGACAGAGGAGCGGCGTTCGCGAGCGAGCGCAAGCGCTTGCGACAGCGG	60	
Db	2509	TCATGCGTTTTTCTTTCAAGCAGAGGGACGTTTCGGGCGAGCGAGCGGAATGGCGTGTCCAAAGGGG	245	
QY	61	AATGGGCGGGCGGTCCGGTGAACGATGCCAAGCCGGTGGCCCAATGTCAAGCGCCCGCGCAT	120	
Db	2449	AATGGAGGCTTTCGGCGGTCAAGATGCCAGGTGATCCGCAAGCTCAACCGCGCGCGCAT	239	
QY	121	GAAGCGCTGCATTCACAGCGGCTCGGCGAGCGGTAAGAACGGCTGACAGCAGCGATATAGCGAG	180	
Db	2389	GAACCGTTTCGATCACTCATGCGCGATGCCGTATGAACCGCTGTGACAGAGGATATAGCGCG	233	
QY	181	GTCGGGCGGGCAAGCCGCGGAACAGCATCCGTTTCAGCAGCCGCGAGGAAGCGGTCCGCATC	240	
Db	2329	GTCGGGCGCGAGCGCGCGGAACAGCATCCGTTCCAAAGGGCGCGAAGAGCGGTCCCGGCG	227	
QY	241	CGCGGATTCGATATGGCCCAAGCGCGCAAGCGCGAACG-----GGCGAGCGCGGTGTCTCAG	294	
Db	2269	CGCCCGGTGATTCGGGTATATCGCGGATGAGCGCGCGACAGCGCGGTGTCGTGCCGGCGGCC	221	
QY	295	GTCGCGCGCGCGATGGCATCCGCGACCTTGCGCGGCGATATAGGCGAGCGAATATCCGGTGAC	354	
Db	2209	GGAACGAGCCCGCACACCAAGTCGCGCACTGTGCGCGCATATAGGCGAGCAATATAGCCGTGAC	215	
QY	355	GGGGTGAACACGCTTGCCCCCAGCGCCAACGGGACACCGCCCCCTTGGCGGTGTGCGCGCA	414	
Db	2149	CGGATGAAGAACCCCGCGCGCATGTCCAAGGGAACAGGCCCCCGCGGTGATGTGCGCCA	209	

QY 415 GAAGCTTATGGCTCATGGGCGACGCGATGGGAGAGATGCCCTTTGCGCGCATCTC 474
 DB 2089 GAAAGCCCGCGATCATGGGCGACGCGATGGGAGATGGCGGCTTCCGCGCGACCTC 2030
 QY 475 CTGCGCGGTCAGCCCGCTGCGCGATATGTCACGACGCTGCGCGACGCGCGCATC 534
 DB 2029 GCGCCCGGTCAGCCCGCTGCGCGATATGTCAGGAGCGCGCGCGACGCGCGCTGC 1970
 QY 535 GTCCAGATGCGCGCGCTGCTGATGAGCGCGATCTCTGATCAGGATGCGGAGTGGAGTGA 594
 DB 1969 GTCCAGATGCGCGCGCTGCTGATGAGCGCGATCTCTGATCAGGATGCGGAGTGGAGTGA 1910
 QY 595 GCGCAGCAGATAGATAGAGCGGATCCCGTTCATCTGCGGAGCGGTCGCTCATGATCAT 654
 DB 1909 GGGCAGCAGATAGATAGAGCGGATCCCGTTCATCTGCGGAGCGGTCGCTCATGATCAT 1850
 QY 655 GCGGCGCTGACGCGCATGGGCGCGCTGCTGCTGATCTGACGCGCGCATATTTCTGAA 714
 DB 1849 GCGGCGCGGCGACGCGCTGCGGCGGCTGCTGCTGATCTGACGCGCGCATATTTCTGAA 1790
 QY 715 ACCCAGCGTCAAGTGGGCGGCTGCGACGCGCGATGCGGATGCGGAGCGGCGCTC 774
 DB 1789 ACCCAGCGTCAAGTGGGCGGCTGCGACGCGCGATGCGGATGCGGAGCGGCGCTC 1730
 QY 775 GATCCGCGAGCGGCTGCGGATGCGGCGGCTGCGGATGCGGATGCGGATGCGGATGCGGAT 834
 DB 1729 GATCCGCGGCTGCGGAGGATGCGGCGGCTGCGGATGCGGATGCGGATGCGGATGCGGAT 1670
 QY 835 CCACCGCAGATGACACCTCTGACGACG-----CCGATCAGCGCGCGCTGATGCA 888
 DB 1669 CCAGCGGATCTGCGCGCGCGCGACCGCGACCGCGATCCGCGCGCGCGCGCGCGCGCA 1610
 QY 889 GCGATAGCTGTCGTCAGGCGCGCGCGATGTCGCGGAAACGCGACCTCTGATCCGTCGA 948
 DB 1609 CCGGTACCGGTGCGCGCGCGCGCGCGATGCGGCGGAAACGCGACCTCTGATCCGTCGA 1550
 QY 949 TTCGCGCGCAGCAATAGGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1008
 DB 1549 GTTGGCG 1490
 QY 1009 GCGAGCAGAGTGTGCTGCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1068
 DB 1489 GCGAGCAGAGTGTGCTGCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1430
 QY 1069 ATCCGCTGCG 1128
 DB 1429 GTTGGCG 1370
 QY 1129 GATCAGCAGATCATGCTCA 1148
 DB 1369 CAGCAGCAGCTCATGCTCA 1350

RESULT 14
 US-09-335-919-7/c
 Sequence 7, Application US/09335919

Patent No. 6150130
 GENERAL INFORMATION:
 APPLICANT: MISAWA, No. 6150130hiko
 APPLICANT: KONDO, Keiji
 APPLICANT: KAJIMURA, Susumu
 APPLICANT: YOKOYAMA, Akihito
 TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
 TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/335,919
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/663,310
 FILING DATE: 23-SEP-1996
 APPLICATION NUMBER: WO PCT/JP94/02220
 FILING DATE: 26-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-235917
 FILING DATE: 05-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 5-348737
 FILING DATE: 27-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 49441/109
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2886 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-09-335-919-7
 Query Match 55.8%; Score 640.8; DB 3; Length 2886;
 Best Local Similarity 73.4%; Pred. No. 6.1e-122;
 Matches 851; Conservative 0; Mismatches 297; Indels 12; Gaps 2;

QY 1 TCATCTCTCTCTCTGACAGAGGCGCGCTTGGGCGAGCGCGCGCTGCGACGCG 60
 DB 2509 TCATCTCTCTCTCTCTGACAGAGGCGCGCTTGGGCGAGCGCGCGCTGCGACGCG 2450
 QY 61 AATGGCGCGCGCTGCGGTCAGATGCGGAAAGCGGTCGCGCATGTCAGGCGCGCGCAT 120
 DB 2449 AATGGAGGCTTGGCGGTCAGATGCGGAAAGCGGTCGCGCATGTCAGGCGCGCGCAT 2390
 QY 121 GAAGCGCTCATCAGCGGCTGCGGCGGCGGAGGCGGTCGAGCGCGCGCGCGCGCGCG 180
 DB 2389 GAAGCGCTCATCAGCGGCTGCGGCGGCGGAGGCGGTCGAGCGCGCGCGCGCGCGCG 2330
 QY 181 GTGCG 240
 DB 2329 GTGCG 2270
 QY 241 GCGCGCATGATGCG 294
 DB 2269 GCG 2210
 QY 295 GTGCG 354
 DB 2209 GAGCAGACCCCGCACACGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2150
 QY 355 GGGGTGAAACAGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 414
 DB 2149 CGGATGAAAGAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2090
 QY 415 GAAGCTTATGGCTCATGGGCGACGCGCGATGCGGCGGATGCGGCGGCGCGCGCATCTC 474
 DB 2089 GAAAGCCCGCGATCATGGGCGACGCGCGATGCGGAGGATGCGGCGGCGCGCGCATCTC 2030
 QY 475 CTGCGCGGTCAGCCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATC 534

```

Db      2029 GGGCCCGGCGCCAGCTGCTGCGGGGCAATGCTGCGGACGCGCGCGCAGCGCTCGTC
Qy      535 GTCCAGATCGCCCGCGTGTGATCGGCTATCTTGATCAGGATCGGGGTGGACTGA
Db      1969 GTCCAGATCGCCCGCGTGTGATCGGCTATCTTGATCAGGATCGGGGTGGACTGA
Qy      595 GGGCAGCAGATGATGAAGCGGTACCCGCTCATCTGCGGACGCTCGCTCCATGATCAT
Db      1909 GGGCAGCAGATGATGAAGCGGTACCCGCTCATCTGCGGACGCTCGCTCCATGATCAT
Qy      655 CGGCGCTTCGACCGCAATGCGGCGCGTGTGATCTGATCTGACGCCCAAGATTTCGAA
Db      1849 CGGCGCGGCGACGCGGTGGGGCGGCTGCTGATCTGACACCCCAATTTCTGAA
Qy      715 ACCCAGCGTCAAGTGGGGGTGTGACAGCGGACCAAGGGGTGATTCAGCAGCAGCGCTC
Db      1789 ACCCAGCGTCAAGTGGGGGTGTGACAGCGGACCAAGGGGTGATTCAGCAGCGCTC
Qy      775 GATCCGCGAGCGGTGCTGATCGCGCGCGGTATCTGACAGCTGCGACATGGGTATT
Db      1729 GATCCGCGGTGCGCAGAGACAGGTGCGCCCTGCGCATCAGCAGGGCGATGCTGTC
Qy      835 CCACCGCAGATGACACCTCTGACGACG-----CCGATCAGCGCGCCGCTCATGCA
Db      1669 CCAGCGGATCTCGCGCGCGCGACCGGACACCGCATCCGCGACGCGCGCTCAGCGA
Qy      889 GCGATAGCTGCTGCTGACAGCGCGCGCAATGTGCGGAAACGCGACTCTGATCCGTCA
Db      1609 CCGTAACTCGGTGCGCGCGCGCGCATGCGGGGAAAGCGCACCCTCTGTCTGCGGCA
Qy      949 TTGCGCGCAGCAATGCGGCGCAGCGCGCGCAGCCATTCGCGCGGAAAGTCCGTCTGTG
Db      1549 GTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      1009 GCAAGCAGCGGTGTGCTGCTGCGAGGGCGCGGACCGCGCTGAGCATGATGAGCGCGC
Db      1489 GCAGGACCGAGTGTGCGCGCTGACAGCTGCTGCGGATGATGATGAGCAGCAGCGCAG
Qy      1069 ATCCGCTGCGGCGTGGCGGACCGGCAAGCGGATGACGCGCAGCGCGCGCGCGCGC
Db      1429 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      - 1129 GATCAGCAGATCATGCTCA 1148
Db      1369 CAGCAGCAGTCTATGGCTCA 1350

```

RESULT 15
US-08-095-726-13/c

```

; Sequence 13, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Profitit, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huei-Chue B
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095, 726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785, 566
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-095-726-13

```

```

Query Match      17.4%; Score 200; DB 1; Length 1235;
Best Local Similarity 50.8%; Pred. No. 5e-32;
Matches 563; Conservative 0; Mismatches 530; Indels 15; Gaps 3;

```

```

Qy      45 CGGCTGCGACAGCGGAATGCGCGCGTCCGCTGACAGTCCGAAGCCGCTCGCCAAATG
Db      1130 CAGACTTTCGCGCGGAACCGGTGCTTCCCTGTAATTCGCGCGCTTATTAAGAGAG
Qy      105 TCAGCGCGCGCGCATAGAGCGCTTCGATCAGCGCTGCGGCGAGCGGTAGAAACCGCTGCA
Db      1070 AGAGCGCAGCGCGCGTAAGCGCTCTACGTGCGCTCCGCGAGCCCATAAAGCCCTGCA
Qy      155 GCAGCGCATAGCAGCGTTCGCGCGCGCGCGCGCGGAAAGAGATCCGCTTCAGCGCGCA
Db      1010 TCACCGCGCAGCGGTCTCTCTCGCGCGCGCGCGCGGAAAGATCCGCTTCAGCGCGCA
Qy      225 GGAAGC---GGTCCGATTCGCGCGCGCGCGCATATGCGCGCGCGCGCGCGCGCGCG
Db      950 AGAATCCCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      282 ACGCGGTGTCAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      890 CGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      342 AATATCCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      830 AATAGCCAGGTGATGAGGAAATAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      402 CGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      770 TCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      462 CGCGCGCGCATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      713 CGCGTTCAAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      522 CAGCGCGCGCGCATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      653 GTTAGGCGATTATCATCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      582 GGGTGGAGCTGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      533 TGTGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      642 CGTTCATGATCATCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      533 CATCATCAGATCGGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy      702 CGAATTTCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db      473 GAAAGAGCTGATAGCGCGGTTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

```

```
QY 762 CGCAGGCGAGCTCGATCCGAGACCGTCCGTCAGCGTCGCGGATATCGTCCAGCGTCG 821
Db 413 CCGCTCCGCGACAGACGCGCTCACCGTTGGCAAGCGCACGCTATTGGGTAAACCTCGC 354
QY 822 CGACATGCGTATTCCACCGAGATCGACACCC-----TGCAGAGCCCGATCAGCG 872
Db 353 TCACCGAAGAGTTTAGCAGATGTTCTCCCGCAGCGCTGATGCAAGGCTCGCAAGG 294
QY 873 CGCCGCTCGATCGAGCCATAGCTGTGTCAAGCGCGCGGATGTGCGGAAACGCGA 932
Db 293 GCTGTAGGTAAATGAGTAGTAGCCGCGCGAGCGACGCGAGATCGGAAACTGCA 234
QY 933 CCTCCGATCCGTCATTCGCGCGAGAGATGCGCGACAGCGCGCCAGCCATTGCGGCG 992
Db 233 CCTCATAGCCCGCGCAGCGGTGCGCACCGCGGCGCAGCGCGTGTGCGCGGAG 174
QY 993 AAGATCCGTGTGTGTCAGAGACAGGTGTGTGTCCAGGCGCGCGACCGCGTCGA 1052
Db 173 TCAGATCGTCTTCATGGAATGACCAAGTATGTTCCCGCGGCTGTCTCCCGGCTCGA 114
QY 1053 GCATCAGATGCGCGCATCCGCTCTGCGGTGCGAAAGCGAGCGGATCAGCGCACCG 1112
Db 113 TCAGCAGCAGGTAACTGCGGCTGAGCGTGTGCGCAGACGCCAGCGATCAGCCGTTGG 54
QY 1113 ACAGCCCCCGCGCCCGGATCAGCAGATC 1140
Db 53 CAGGCGCGCGCGGACTTAAATCAGATC 26
```

Search completed: June 22, 2005, 01:09:49
Job time : 312.581 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 15:45:26 ; Search time 751.928 Seconds
(without alignments)
9045.789 Million cell updates/sec

Title: US-10-695-980-1_COPY_5794_6942

Perfect score: 1149
Sequence: 1 tcatgctctctctcgcagca.....atcagcagatcatgctcat 1149

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 295987067 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1149	100.0	8625	2 AAT45143	Aat45143 Flavobact
2	1149	100.0	8625	2 AAV40146	Aav40146 Flavobact
3	1147.4	99.9	11233	2 AAV40151	Aav40151 DNA sequ
4	640.8	55.8	1161	2 AAG99489	Aag99489 3 hydroxy
5	640.8	55.8	1161	2 AAV84076	Aav84076 Carotenol
6	640.8	55.8	1161	2 AAX19116	Aax19116 Agrobacte
7	395.4	34.4	1966	11 ADQ96933	Adq96933 CrtrCrtY
8	394	34.3	1944	11 ADQ96933	Adq96933 CrtrCrtY
9	236.4	20.6	1185	10 ADC99013	Adc99013 Bradyrhiz
10	213.4	18.6	12783	12 ADM98599	Adm98599 Geranylge
11	209.8	18.3	1149	6 AAD35511	Aad35511 Pantoea s
12	209.8	18.3	1149	8 ABT14191	Abt14191 Pantoea s
13	209.8	18.3	1149	8 ACC44761	Acc44761 Pantoea s
14	209.8	18.3	1149	12 ADQ14628	Adq14628 Pantoea s
15	209.8	18.3	1149	12 ADQ14628	Adq14628 Pantoea s
16	209.8	18.3	1149	12 ADQ48638	Adq48638 Pantoea s
17	209.8	18.3	1149	12 ADQ94455	Adq94455 Crty codi
18	209.8	18.3	8609	12 ADQ14666	Adq14666 Reporter
19	209.8	18.3	8609	12 ADQ26394	Adq26394 Carotenol
20	209.8	18.3	8609	12 ADQ77262	Adq77262 Nucleotid

C 21	209.8	18.3	8609	12	ADQ48673	Adq48673 Plasmid p
C 22	209.8	18.3	8609	12	ADQ94478	Adq94478 Plasmid p
C 23	200	17.4	1235	2	AAT40795	Aat40795 Lycopen
C 24	200	17.4	1235	2	AAT91547	Aat91547 Erwinia h
C 25	198.4	17.3	1234	2	AAQ13722	Aaq13722 Lycopen
C 26	198.4	17.3	1234	2	AAQ13723	Aaq13723 Lycopen
C 27	198.4	17.3	1235	2	AAT40796	Aat40796 Recombina
C 28	198.4	17.3	1235	2	AAT91548	Aat91548 Genetica
C 29	174.8	15.2	1149	2	AAQ06295	Aaq06295 Sequence
C 30	174.8	15.2	1149	2	AAV84083	Aav84083 Carotenol
C 31	174.8	15.2	1149	2	AAK19120	Aak19120 Erwinia u
C 32	174.8	15.2	6524	12	ADP74122	Adp74122 Pantoea a
C 33	174.8	15.2	6918	2	AAQ06299	Aaq06299 Sequence
C 34	174.8	15.2	7494	8	ABZ69177	Abz69177 Vector co
C 35	174.8	15.2	8547	8	ABZ69178	Abz69178 Vector co
C 36	150.8	13.1	1167	2	AAV73182	Aav73182 C. utrilis
C 37	138.8	12.1	1631	12	ADQ61158	Adq61158 Alcaligen
C 38	138.8	12.1	1631	12	ADQ61075	Adq61075 Alcaligen
C 39	138.8	12.1	1631	13	ADQ38244	Adq38244 Alcaligen
C 40	138.8	12.1	1631	13	ADQ03859	Adq03859 Alcaligen
C 41	138.8	12.1	1631	13	ADR03939	Adr03939 Alcaligen
C 42	99.6	8.7	1176	10	ACF70994	Act70994 Photorhab
C 43	99.6	8.7	110000	10	ACF67367	Act67367 Photorhab
C 44	99.6	8.7	110000	10	ACF65388	Act65388 Photorhab
C 45	99.6	8.7	110000	10	ACF65388_07	Act65388_07 Photorhab

ALIGNMENTS

RESULT 1	
ID AAT45143	standard; DNA; 8625 BP.
XX AAT45143;	
AC	
XX	
DT 17-OCT-2003	(revised)
DT 08-MAR-1997	(first entry)
XX	
DB Flavobacterium carotenoid biosynthesis cluster DNA.	
XX	
KW Carotenoid; lycopen; beta-carotene; echinenone; canthaxanthin;	
KW zeaxanthin; adonixanthin; astaxanthin; crtE;	
KW geranylgeranyl pyrophosphate synthase; GGPP synthase; crtB;	
KW prephytyene synthase; phytyene synthase; crtI; phytyene desaturase; crtY;	
KW lycopen cyclase; crtZ; beta-carotene hydroxylase; ds.	
XX	
OS Flavobacterium sp. ATCC 21588; WT (ATCC 21588).	
XX	
FH	
FT CDS	location/Qualifiers
FT	2..1168
FT	/*tag= a
FT	/label= ORF-5
FT	1180..2355
FT	/*tag= b
FT	/label= ORF-1
FT	2521..3408
FT	/*tag= c
FT	/label= crtB
FT	/*tag= C
FT	/product= "GGPP synthase"
FT	complement(3405..4316)
FT	/*tag= d
FT	/label= crtB
FT	/*tag= C
FT	/product= "phytyene synthase"
FT	complement(4313..5797)
FT	/*tag= e
FT	/label= crtI
FT	/*tag= C
FT	/product= "phytyene desaturase"
FT	complement(5794..6942)
FT	/*tag= f
FT	/label= crtY
FT	/*tag= C
FT	/product= "lycopen cyclase"
FT	complement(6939..7448)
FT	CDS

```
FT      /*tag= g
FT      /label= Cry2
FT      /product= "beta-carotene hydroxylase"
FT      /complement(7767. .8315)
FT      CDS
FT      /tag= h
FT      /label= ORF-16
FT      misc_difference 8348. .8349
FT      /tag= 1
FT      /note= "bases 8348-8349 are given as nm in the
FT      specification"
FT      misc_difference 8539. .8540
FT      /tag= 1
FT      /note= "bases 8539-8540 are given as nm in the
FT      specification"
FT      misc_difference 8590
FT      /tag= 1
FT      /note= "base 8590 is given as n in the specification"
FT      misc_difference 8592
FT      /tag= m
FT      /note= "base 8592 is given as n in the specification"
FT      misc_difference 8602. .8604
FT      /tag= n
FT      /note= "bases 8602-8604 are given as nm in the
FT      specification"
FT      EP747483-A2.
FT      11-DEC-1996.
FT      29-MAY-1996; 96EP-00108556.
FT      09-JUN-1995; 95EP-00108888.
FT      (HOFF ) HOFFMANN LA ROCHE & CO AG F.
FT      Hohmann H, Pasamontes L, Tessier M, Van Loon A;
FT      WPI, 1997-023160/03.
FT      P-PDB; AAM06513, AAM06514, AAM06515, AAM00871, AAM06517,
FT      AAM06518, AAM06519.
FT      Probacterium gene sequences encoding carotenoid biosynthesis enzymes -
FT      for the production of carotenoid(s), useful in foods and animal feeds.
FT      Example 2; Fig 7; 80pp; English.
FT      Genomic DNA (AAT45143) of Flavobacterium sp. R1534 includes genes of the
FT      carotenoid biosynthesis pathway. The sequence was deduced from inserts of
FT      6 clones obtd. from genomic libraries e.g. by PCR amplification (see also
FT      CC AAT5144-45) and use of partial clones to screen the library. The
FT      CC identities of the gene products (see also AAM06513-19 and AAM00871) were
FT      CC detd. by examining carotenoid accumulation in E. coli hosts transformed
FT      CC with deleted variants of the gene cluster. The isolated genes can be used
FT      CC in different combinations to produce carotenoids in transformed host
FT      CC cells. (Updated on 17-OCT-2003 to standardise OS field)
XX      SQ
XX      Sequence 8625 BP, 1458 A, 2898 C, 2964 G, 1295 T, 0 U, 10 Other;
SQ      Query Match      100.0%; Score 1149; DB 2; Length 8625;
SQ      Best Local Similarity 100.0%; Pred. No. 1.3e-206;
SQ      Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 TCATGCTCTCTCTGAGCAGGAGGCGTTTCGAGGAGCAGCGCAGCGCTTGCAGACGCG 60
DB      TATATGCTCTCTCTGAGCAGGAGGCGTTTCGAGGAGCAGCGCAGCGCTTGCAGACGCG 5853
QY      61 AATGGCGGCGGCTCCGCTGACGATGCGAGCGCGTGGCGCATGTCTAGCGCGCCGCGCAT 120
DB      TATGCTCTCTCTGAGCAGGAGGCGTTTCGAGGAGCAGCGCAGCGCTTGCAGACGCG 5853
QY      5854 AATGGCGGCGGCTCCGCTGACGATGCGAGCGCGTGGCGCATGTCTAGCGCGCCGCGCAT 5913
DB      TATGCTCTCTCTGAGCAGGAGGCGTTTCGAGGAGCAGCGCAGCGCTTGCAGACGCG 5853
```

```
QY      121 GAAGCGCTCATGTCAGCGGCTGCGGAGCGGCTGAGAACCGCTTGCAGCAGCGCATGACGACG 180
DB      GAAAGCGCTCATGTCAGCGGCTGCGGAGCGGCTGAGAACCGCTTGCAGCAGCGCATGACGACG 5973
QY      181 GTTCGGCGGCGGAGCGCGGAGAACGATTCGGTTGACGACCGCGGAGAACGGGTGGCGATC 240
DB      GTTCGGCGGCGGAGCGCGGAGAACGATTCGGTTGACGACCGCGGAGAACGGGTGGCGATC 6033
QY      241 CGCGGATCATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATTCGATGAGTTCGCG 300
DB      CGCGGATCATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATTCGATGAGTTCGCG 6093
QY      301 CGCGGATCATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATTCGATGAGTTCGCG 360
DB      CGCGGATCATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATTCGATGAGTTCGCG 6153
QY      361 GAACAGCGGCGGCGGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATTCGATGAGTTCGCG 420
DB      GAACAGCGGCGGCGGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATTCGATGAGTTCGCG 6213
QY      421 TATGCGCTCATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 480
DB      TATGCGCTCATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6273
QY      481 GGTCCAGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 540
DB      GGTCCAGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6333
QY      541 ATTCGCGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 600
DB      ATTCGCGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6393
QY      601 CAGATGATGAGAGCGGATCCGCTCATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 660
DB      CAGATGATGAGAGCGGATCCGCTCATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6453
QY      661 CTGACGCGCATGAGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 720
DB      CTGACGCGCATGAGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6513
QY      721 GGTGAGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 780
DB      GGTGAGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6573
QY      781 CGAGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 840
DB      CGAGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6633
QY      841 CAGATGACACCGCTGAGCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 900
DB      CAGATGACACCGCTGAGCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6693
QY      901 CGTCAAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 960
DB      CGTCAAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6753
QY      961 AATGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 1020
DB      AATGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6813
QY      1021 GTGCTGTCGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 1080
DB      GTGCTGTCGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6873
QY      1081 GTTCGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 1140
DB      GTTCGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGATTCGATGAGTTCGCG 6933
QY      1141 ATGCTCAT 1149
DB      ATGCTCAT 6934
```

[illegible]

Accession	Gene	Strain	Position	Sequence	Length
Db			6094	CGCCGCGATGGCATTCGGGACCTTGGGCGCATAGGGGAGCGGAAATTCGGGTGACGGGGTGG	6153
Qy			361	GAAAGAGCCCTTGGCCCCCAGGCCCAACCGGACCCGCCCTGGCGGTGGTTCGGCCAGAAAGCC	420
Db			6154	GAAAGAGCCCTTGGCCCCCAGGCCCAACCGGACCCGCCCTGGCGGTGGTTCGGCCAGAAAGCC	6213
Qy			421	TATGGCGGTCAATGGGCGCAGCGCGATGGGGCAGGATGCGCTTTTGGCGCGCATCTCTGCCC	480
Db			6214	TATGGCGGTCAATGGGCGCAGCGCGATGGGGCAGGATGCGCTTTTGGCGCGCATCTCTGCCC	6273
Qy			481	GGTCCAGCCCCCGCTGGGCGGCGCATAGTCCAGGAGAGCGGCGGCGGCGGCGCATGCTGCAG	540
Db			6274	GGTCCAGCCCCCGCTGGGCGGCGCATAGTCCAGGAGAGCGGCGGCGGCGGCGCATGCTGCAG	6333
Qy			541	ATCCGCGCGCGTCCGCTGTAGCGCGGTATCTCTCATCAGATGCGGGTGGGACTGAAGGGCAG	600
Db			6334	ATCCGCGCGCGTCCGCTGTAGCGCGGTATCTCTCATCAGATGCGGGTGGGACTGAAGGGCAG	6393
Qy			601	CAGATAGATGAAGCGGGTACCCTGTCATCTTGGGAGCGGTCGCTCATGATCATCGGGCG	660
Db			6394	CAGATAGATGAAGCGGGTACCCTGTCATCTTGGGAGCGGTCGCTCATGATCATCGGGCG	6453
Qy			661	CTCCAGCGCGAATGGGGGGGGTCCGCTCGATCTCGAGCGCCACGAAATTTCTGGAAACCCAC	720
Db			6454	CTCCAGCGCGAATGGGGGGGGTCCGCTCGATCTCGAGCGCCACGAAATTTCTGGAAACCCAC	6513
Qy			721	GGTCAGGTGCGGGGTCTCGACGCGACGCGGCGCTGCATCAGCGAGCGAGCTTCGATCCG	780
Db			6514	GGTCAGGTGCGGGGTCTCGACGCGACGCGGCGCTGCATCAGCGAGCGAGCTTCGATCCG	6573
Qy			781	CGAGCCGCTCCGTCAAGCGCTCGCGCGGTATGTCAGCGGTGCGCATGCGTATTTCACCG	840
Db			6574	CGAGCCGCTCCGTCAAGCGCTCGCGCGGTATGTCAGCGGTGCGCATGCGTATTTCACCG	6633
Qy			841	CAGATTCGACACCTTCGACAGCGCCCGATCAGCGCGCGCGCGCTCGATTCAGAGCCATAGCCGT	900
Db			6634	CAGATTCGACACCTTCGACAGCGCCCGATCAGCGCGCGCGCGCTCGATTCAGAGCCATAGCCGT	6693
Qy			901	CGTCAGGCGGGCGCGAATGTCGGGAAACGCGACCTCTCGATCCGTCATTCGCGCGCAGC	960
Db			6694	CGTCAGGCGGGCGCGAATGTCGGGAAACGCGACCTCTCGATCCGTCATTCGCGCGCAGC	6753
Qy			961	AATGGGCGACAGGCGCGCGCACCATTCGCGGAGAAAGATCCGCTGTGTCGGCAGACAGGT	1020
Db			6754	AATGGGCGACAGGCGCGCGCACCATTCGCGGAGAAAGATCCGCTGTGTCGGCAGACAGGT	6813
Qy			1021	GTGCTGGTTCGAGGGGGCGCGACCGCGCGTTCGAGATCAAGATGGCGCCATTCGGGTCTGGC	1080
Db			6814	GTGCTGGTTCGAGGGGGCGCGACCGCGCGTTCGAGATCAAGATGGCGCCATTCGGGTCTGGC	6873
Qy			1081	GTCCGGAACGGCAAGCGGATCAGCGCACCGGACGAGCGCGCGCGCGCGATCAGCAGATC	1140
Db			6874	GTCCGGAACGGCAAGCGGATCAGCGCACCGGACGAGCGCGCGCGCGCGATCAGCAGATC	6933
Qy			1141	ATGGCTCAT 1149	
Db			6934	ATGGCTCAT 6942	
RESULT 3					
AAV40151					
ID AAV40151 standard; DNA; 11233 BP.					
XX AC AAV40151;					
XX AC 10-AUG-1999 (first entry)					
XX AC 10-AUG-1999 (first entry)					
DE DNA sequence of plasmid pZea4.					
KM Carotenoid; pigment; canthaxanthin; R134; crtB; prephycoene synthase;					
KM crtI; phycoene desaturase; crtY; lycopene cyclase; crtW396; pZea4;					
KM beta-carotene beta-oxygenase; food product; fermentation; db.					

XX Synthetic.
OS JP10155497-A.
PN 16-JUN-1998.
XX 02-DEC-1997; 97JP-00348653.
XX 02-DEC-1996; 96EP-00810839.
XX (HOPF) HOPFMANN LA ROCHE & CO AG F.
XX WPI; 1998-391048/34.
DR Preparation of carotenoid - comprises fermentation with transformed cell.
XX Example 2; Fig 42-53; 80pp; Japanese.
PS The invention describes the preparation of carotenoid pigments e.g.
XX canthaxanthins using a cell transformed by a vector having DNA sequences
CC (a) to (e) or substantially homologous sequences. (a) a DNA sequence
CC (crtB) coding GSP synthase of Flavobacterium sp. R1534; (b) a DNA
CC sequence (crtB) coding prephytyene synthase of Flavobacterium sp. R1534;
CC (c) a DNA sequence (crtI) coding phytyene desaturase of Flavobacterium
CC sp. R1534; (d) a DNA sequence (crtY) coding lycopene cyclase of
CC Flavobacterium sp. R1534; and (e) a DNA sequence (crtW336) coding beta-
CC carotene beta-oxygenase of a microbe E-396 (FERM BP-4283). The carotenoid
CC or a carotenoid mixture can also be used in preparation of food products.
CC The method is an improved method of fermentation for carotenoid
CC production
SQ Sequence 1123 BP; 2142 A; 3525 C; 3605 G; 1960 T; 0 U; 1 Other;

Query Match 99.9%; Score 1147.4; DB 2; Length 11233;
Best Local Similarity 99.9%; Pred. No. 2.6e-206;
Matches 1148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCATGCTCTCTCTGAGAGAGGGGGGCTTCGGGACAGCGGACCGCTTGGACAGCGG 60
DB 6476 TCACTGCTCTCTCTGAGAGAGGGGGGCTTCGGGACAGCGGACCGCTTGGACAGCGG 6535
QY 61 AATGGGCGGGCGGTCCGGTACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCCGGCAT 120
DB 6536 AATGGGCGGGCGGTCCGGTACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCCGGCAT 6595
QY 121 GAAGCGCTGATCAGCGGCTGCGGCAAGCGGCTGGAACCCCTGACAGCGGATAGCGAGC 180
DB 6596 GAAGCGCTGATCAGCGGCTGCGGCAAGCGGCTGGAACCCCTGACAGCGGATAGCGAGC 6655
QY 181 GTCGGGCGGGACGCGGGAACGATCCGGTTCAGCAGCCGACAGGAACGGATCCGATC 240
DB 6656 GTCGGGCGGGACGCGGGAACGATCCGGTTCAGCAGCCGACAGGAACGGATCCGATC 6715
QY 241 CCGCGATGATGCGCCAGCCGCGACCGCGGACCGCGGAGCGCGGTCGTCAAGTCCG 300
DB 6716 CCGCGATGATGCGCCAGCCGCGACCGCGGAGCGCGGTCGTCAAGTCCG 6775
QY 301 CCGCGGATGCGATCCGCGAAGCTGCGGCGGATAGGGACGGAATATCCGGTGAACGGGTC 360
DB 6776 CCGCGGATGCGATCCGCGAAGCTGCGGCGGATAGGGACGGAATATCCGGTGAACGGGTC 6835
QY 361 GAACAGCCCTGCGCCCGACGCGGCAACCGCGCCCTGCGGTCGGCGCAGAGGCC 420
DB 6836 GAACAGCCCTGCGCCCGACGCGGCAACCGCGCCCTGCGGTCGGCGCAGAGGCC 6895
QY 421 TATGGCGTATGCGGCGGCGGATGCGGAGATGCGGTCCTTTCGCGCGCATCTCTGCC 480
DB 6896 TATGGCGTATGCGGCGGCGGATGCGGAGATGCGGTCCTTTCGCGCGCATCTCTGCC 6955
QY 481 GGTCCAGCCCGCTGCGGCGGATATGTCAGCGAGCGCTTCCGCGCAGCGCCCATCGTCCAG 540
DB 6956 GGTCCAGCCCGCTGCGGCGGATATGTCAGCGAGCGCTTCCGCGCAGCGCCCATCGTCCAG 7015

QY 541 ATCCGCCGCTGCTGTAGCCGCTATCTTCATCAGATGCGGGTGGAGCTGAAGGGCAG 600
DB 7016 ATCCGCCGCTGCTGTAGCCGCTATCTTCATCAGATGCGGGTGGAGCTGAAGGGCAG 7075
QY 601 CAGATGATGAAGCGGTCATCCGTCATTCGGGAAACGGTCCGCTCATGATCATCGGCG 660
DB 7076 CAGATGATGAAGCGGTCATCCGTCATTCGGGAAACGGTCCGCTCATGATCATCGGCG 7135
QY 661 CTCGACGCGATGCGGGGGGTGGTCTGATCTCGACGCGCCAGATTTCTGGAACCCAC 720
DB 7136 CTCGACGCGATGCGGGGGGTGGTCTGATCTCGACGCGCCAGATTTCTGGAACCCAC 7195
QY 721 GGTCAAGTGCAGGGGTCTCGACGCGCACACGCGGCGTGCATCGACGAGCGCTCGATCCG 780
DB 7196 GGTCAAGTGCAGGGGTCTCGACGCGCACACGCGGCGTGCATCGACGAGCGCTCGATCCG 7255
QY 781 CGAGCCGTCCTGACGCTGCGGCGGTCATCTCGACGCGTCGAGATCGCATTCACCG 840
DB 7256 CGAGCCGTCCTGACGCTGCGGCGGTCATCTCGACGCGTCGAGATCGCATTCACCG 7315
QY 841 CAGATCGACACCTGCGACAGCCGATGAGGCGCGCGCTCGATCGAGCCATAGCCTGT 900
DB 7316 CAGATCGACACCTGCGACAGCCGATGAGGCGCGCGCTCGATCGAGCCATAGCCTGT 7375
QY 901 CGTCAAGCGGCGCGAATGTCGGGAAACGCGACCTCTGATCCGTCATTGCGCGCGACG 960
DB 7376 CGTCAAGCGGCGCGAATGTCGGGAAACGCGACCTCTGATCCGTCATTGCGCGCGACG 7435
QY 961 AATGGGCGACAGGCGCGCCGACCATTTCCGGGCGAAAGATCCGTCGTGCGCAGACAGT 1020
DB 7436 AATGGGCGACAGGCGCGCCGACCATTTCCGGGCGAAAGATCCGTCGTGCGCAGACAGT 7495
QY 1021 GTGCTGTCCGAGGGGCGGACCGCGCGTCGACGATCGAGTGGCGATCCGGCTGCG 1080
DB 7496 GTGCTGTCCGAGGGGCGGACCGCGCGTCGACGATCGAGTGGCGATCCGGCTGCG 7555
QY 1081 GTCGGAACGCGGACCGGATCGAGCGCACCGGACAGCCCGCGCGGATCGAGCATC 1140
DB 7556 GTCGGAACGCGGACCGGATCGAGCGCACCGGACAGCCCGCGCGGATCGAGCATC 7615
QY 1141 ATGCTCAT 1149
DB 7616 ATGCTCAT 7624

RESULT 4
AA09489/c
ID AA09489 standard; DNA; 1161 BP.

XX AA09489;
AC 27-AUG-2003 (revised)
DT 28-FEB-1996 (first entry)
XX 3 hydroxy-beta-ionone ring methylene to keto converting peptide DNA.
XX Xanthophyll; astaxanthine; methylene; keto group; conversion;
XX 3-hydroxy-beta-ionone ring;
XX Agrobacterium aurantiacum.
XX
XX
XX Key Location/Qualifiers
FT 1. 1161
FT /*tag= a
FT /product= "beta-ionone ring 4-methylene_to_4-keto
FT _group_converting_peptide."
PN MO9518220-A1.
XX 06-JUL-1995.
PD 26-DEC-1994; 94WO-JP002220.
PF

XX	New carotenoid glucoside(s) - used as food additives.
PT	
XX	
PS	Disclosure; Page 12-13; 26pp; Japanese.
XX	
CC	The present sequence encodes a protein involved in carotenoid
CC	biosynthesis. The specification describes ataxanthin diglucosides and
CC	adonixanthin-3'-glucosides. The specification also describes a method for
CC	the preparation of a carotenoid glycoside, in which all, or part of,
CC	carotenoid biosynthesis genes <i>crte</i> , <i>crtb</i> , <i>crtl</i> , <i>crty</i> , <i>crtz</i> , <i>crx</i> or <i>crw</i>
CC	are introduced to a microbe or plant and expressed. The transformed
CC	organism is cultured and ataxanthin diglucosides, adonixanthin-3'-
CC	glucosides, and/or ataxanthin monoglucoside are collected. The
CC	carotenoid glucosides are used as food additives
XX	
SQ	Sequence 1161 BP; 148 A; 417 C; 410 G; 186 T; 0 U; 0 Other;

Query Match	55.8%;	Score 640.8;	DB 2;	Length 1161;
Best Local Similarity	73.4%;	Pred. No. 3.2e-111;		
Matches 851; Conservative	0;	Mismatches 297;	Indels 12;	Gaps 2;

QY	1	TCATGCTCTCTCTCAGACAGGAGGGCGCTTCGGGACAGGACCCACGCGCTGCGACGACGCG	60
Db	1161	TCATGCTCTCTCTCAGACAGGAGGGAGCTTCGGGACAGGACCCGATGCGCTGCGACAGGAG	110
QY	61	AATGGGCGGGCGTCCGGTGACGATCGAAGCCGGTCGCGCAATGTCAGGCGCCCGGCATA	120
Db	1101	AATGGAGAGGCTTGGCCGGTCAAGATCGGACGTGATCCGCAACGCTACGCCCGCGGCATA	104
QY	121	GAAAGCGCTCGATCAACGGCGTGGGAGAGCGGTGAGAACCGCTGACAGAGCGCATTAAGAGG	180
Db	1041	GAACTGTCGATCAAGTCATGCGAGGATGGGATGGATGAACCGCTGACAGAGGCTTAAGGCGG	982
QY	181	GTCGGGCGGGGACGCGCGGAAACAGCATCCGGTTTCAGACAGCCGACAGAAACCGGTGCGATC	240
Db	981	GTCGGGCGCGCAGCCGCGGAAACGATCCGGTTCAAAAGACCGCAGAAACCGGTGCGGCG	922
QY	241	CGCGCGATCGATGAGCCCAAGCCGCGACACCGCGGACG-----GGCGAGCGCGATGTCAG	294
Db	921	CGCCCGGTGATCGCGTATCGCGATGCGCGATGGCGCGCAGCGCGTGGTCCGGGCGGCGCC	862
QY	295	GTCGCGCGCGCGGATGCGCATCCGCGACCTGCGCGGCGATAGGGGACGCAATATCCGTGAC	354
Db	861	GGAACAGACCCCGCACACGATCGCGCACCTGTCCGATAGGGGACAGGAATACCCGCTGAC	802
QY	355	GGGGGTGGAACAGCCCTGACCCCGACGCGCAACCGGACCCCGCTGGCGGTGGCGCGCA	414
Db	801	CGGATGAAGAAACCCCCGCGCGCATGTCACAGGAAACAGGCCCGCGCGTGAATCGGCCCA	742
QY	415	GAAAGCTATGCGGTCATAGGCGCAGGCGCATAGGACAGATGCCCTTTCGCGCGCATCTC	474
Db	741	GAAAGCCCGCGCATATAGGCGCAGGCGCATAGGAAAGATGCCGCGTTCGCGCGGACCTC	682
QY	475	CTGCCCCGTCCAGCCCCCGCTTGCGCGCATATGTCACAGCGCCTTGCGCGCAGCGCGCATC	534
Db	681	GAGCCCCGTCCAGCCCCCTGTGCGCGGACATAGTCGATGAGGACGCGCGCGCACGCGCTGTC	622
QY	535	GTCACGATCGCGCGCGCTGCTAGCGGGGTATCCTCGATCAGAGATGCGGGTGGACTGAA	594
Db	621	GTCACGATCGCGCGCATCGAATAGGGCGGTGTCCTCGATCAGAGATGCGGGTGGAGAGAA	562
QY	595	GCGCAGCAGATATGATGAAGCGGTACCCGTCATCTGCGGAAACGGTCGCTCCATGATCAT	654
Db	561	GCGCAGCAGATATGATGAAGCGGTACCCGTCCTGCTGCGGAGCGGTGCGCTCCATGATCAT	502
QY	655	CGGCGCTTCGACGCGCATGGGGGGCGTCGCTCTTCGATCTTCAGCGCCACGAATTTCTGAA	714
Db	501	CGGCGCGGCGACGCGGTGGGGCGCGTCCGCTCTCGATCTTCAGACCCACGAATTTCTGAA	442
QY	715	AACCAACGTCAGGTCGGGGGTTCAGACGCGACACCAAGGGGTGATACGACAGCAGCCTC	774
Db	441	AACCAACGTCAGATCCGCGAGCGCTGCGCGCCCGCGCTCCAGACCGCGCCCGCTC	382

QY 775 GATCCGACAGCCGTCCTCCGACAGCGTCGACCCCGGTATCGTCAGCGCTCGGAATGCGTATT 834

Db 381 GATCCGCGGTCCGACAGGACAGCGGTGTGCCCCCTGGCGCATCCAGAGGGGAGATGTCGCTGC 322

QY 835 CCACCGCAGATCGACACCTTGCAGACG-----CCGATCAGCGCGCCCGCTCGATGGA 888

Db 321 CCACGCGGATCTCGGCGGCCGACCCGAGCCACCGCATCCCGCCAGCGCGCCCGTCCAGCGGA 262

QY 889 GCCATAGCGCTTCTGTCAGAGCGGCGCGGATGTGTGGGAAAACCGCATCTCTGATCCGTGCA 948

Db 261 CCGGTAAACCGGTGGCGCAACGCCGCCCGGGCATGGCGGGGAAAAGCGCATCTTCTGTGTCGGGCA 202

QY 949 TTGCGCGCGGAGAAATGGGCGGACAGGCGCGGCCAGCCATTCCGGGCGGAAAGATCCGTCGTG 1008

Db 201 GTTGGCGCGGCGGACAGGGGCTTCAAGCCGCGGACAGCCAGTCCTGGCGACAGGTCCGGGTCGTG 142

QY 1009 GCAGAGCCAGGATGTGCTGTCGTCCGAGGGGCGGACCGCGCGTCGTGAGCATCAGATCGGCGC 1066

Db 141 GCAAGGACCAAGGTGTGGCGCGTCTGAACGATCTTGCGGCATGTGTCAGCAGCAGGACCGCGCAG 82

QY 1069 ATCCGATTTGCGGTTGCGGAAACGGGCAAGCGCGGATCAGCGCACCGGACAGCCCTCGCCCGCG 1122

Db 81 GTCCGGCGCGCCCGCGCGCAGCGCCGACAGGGCGATCAGCCCGTTGGCGAAGGCCGCGCCCTGCGC 22

QY 1129 GATCAGCAGATCAGGCTCA 1148

Db 21 CAGCAGCAGCTCATAGGCTCA 2

RESULT 6
AA19116/c
ID AA19116 standard; DNA; 1161 BP.

DT 14-MAY-1999 (first entry)

DE Agrobacterium aurantiacum crty encoding DNA

KW Beta-carotene hydroxylase; crty; crtb; crtb; crti; xanthophyll;

KW metabolite; ss.

OS *Agrobacterium aurantiacum.*

PN JP11046770-A.

PD 23-FEB-1999.

PF 07-AUG-1997; 97JP-00213648.

PR 07-AUG-1997; 97JP-00213

PA (KIRI) KIRIN BREWERY KK.

DR WPI; 1999-208113/18.

XX
XX

PT their metabolites.

PS Disclosure; Page 9-11; 17pp; Japanese.

CC The present invention describes beta-carotin hydroxylase. Beta-carotene

CC present sequence represents an *Agrobacterium aurantiacum* crty DNA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match	55.8%;	Score 640.8;	DB 2;	Length 1161;
Best Local Similarity	73.4%;	Pred. No. 3.2e-111;		
Matches 851; Conservative	0;	Mismatches 297;	Indels 12;	Gaps 2

```

QY 1 TCATGCTCTCTCTGAGCAGAGGGGGGCTTCGGGCAAGCAGCGACGGCTTCGACAGCGG 60
DB 1161 TCAATGGTTTCTCTTCAAGAGGGGAGCTTCGGGCAAGCGGATGCGGTGCAAGGGG 1102
QY 61 AATGGGCGGGCGCTCGGTGACATGCGAAGCCGGTGGCCATGTCAAGGCGCCGCAATA 120
DB 1101 AATGGAGGGCTTCCCGCTCAAGATGCGCAGCTGATCCGCAACGCTCAGCGCGCCGCAATA 1042
QY 121 GAAGCGCTGATGAGGCGGCTGGGCGAGGGGTGAACCGCTGACAGAGCGATAGGAGCG 180
DB 1041 GAACCGTGTGATGATGTCATGCGGCAAGCGGTAGAAACCGTGAAGAGGGATAGGCGCG 982
QY 181 GTCGGGCGGGCAGCGCGCGAAGCAGATCCGGTTCAAGCAGCGCGAGAGAGCGGTCCGATC 240
DB 981 GTCGGGCGGGCAGCGCGCGAAGCAGATCCGGTTCAAAAGCGCGAAGAGCGGTCCGCGCG 922
QY 241 CGGCGATGATGAGCGCCAGCGCGCAGCGCGCGACG-----GGCGAAGCGGTGTGAG 294
DB 921 CGCGCGGTGATGCGCTTAATCGCGATGAGCGCGCGCGCGCGCGTGGTCCGCGCGCGCG 862
QY 295 GTCGGCGCGCGGATGAGCATCCGCGCACTCCGCGGATAGGCGAGAGGATATCCGGTGAC 354
DB 861 GAGCAGACCCCGCACCACTGCGCACCTGTGCGGATAGGCGAGGAAATAGCGGTGAC 802
QY 355 GGAGGTGAACAGCGCTGCGCCCGCACCGCGCAACCGCGCACCGCGCGTGGTGTGCGCGCA 414
DB 801 CGGATGAAGAAACCCCGCGCGCGAGTCCACGAGGAAAGCGCGCGCGCGGTGATCGCGCCA 742
QY 415 GAAGCTTATGCGGTCAATGAGGCGCGAGCGGATGAGGCGAGATGCCCTTTGCGCGCGCATCTC 474
DB 741 GAAGCGCGCGCATCAATGAGGCGCGAGCGGATGAGGAGATGCGCGTTCGCGCGCGAGCTTC 682
QY 475 CTGCGCGGTGCGAGCGCGCGCTGCGCGCATATGTCACGAGCGCTGGCGCGCGCGCGCATC 534
DB 681 GCGCGCGGTGCGAGCGCGCTGCTGAGGCGGATGCTGTGAGAGCGCGCGCGCGCGCTGCTC 622
QY 535 GTCAGATGCGCGCGCTGCTGAGGCGCGCATCTCTGATCAGATCGAGATCGAGGTGAGACTGA 594
DB 621 GTCAGATGCGCGCGCATCGGAATAGCGCGGTCTGATCAGATCGAGATCGAGGTGAGAGAA 562
QY 595 GGGCAGCAGATGATGAGAGCGGTACCGGTCTCATCTGCGGAGCGGTGCGGTCAATGATCAT 654
DB 561 GGGCAGCAGATGATGAGAGCGGTACCGGTCTCTGAGTGAAGGTGCGGTGCGGTCAATGATCAT 502
QY 655 CGGCGCTCTGAGCGCCATGAGGGGGCGTCTGCTGATCTGACGCGCCCAAGATTTCTGGA 714
DB 501 CGGCGGGGCGACGCGCTGAGGGGGGTGCTGCTGATCTGACACCCACAAATTTCTGGA 442
QY 715 ACCCAGCTCAGGTGCGGGGTCTGACGCGACACGAGGCGGTGATCAGCGAGCGAGCTTC 774
DB 441 ACCCAGCTCAGGTGCGGGGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 382
QY 775 GATCCGCGAGCGGTCTGCTGAGCGTCTGCGCGCGGTATGTCAGCGGTGCGCATATGCTATT 834
DB 381 GATCCGCGGTGCGGAGACAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 322
QY 835 CCAAGCGCAGATGACACCTCTGACAGAC-----CGATGAGCGCGCGCGCGCGCGCGTCA 888
DB 321 CCAAGCGCAGATCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 262
QY 889 GGCATGAGCTGTGATGAGGCGGCGGAGATGATGCGGAAACGAGACTCTCGATTCGATCA 948
DB 261 CCGGTACCGGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 202
QY 949 TTGCGCGCAGCAATGAGGCGGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1008
DB 201 GTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 142
QY 1009 GAGAGCAGAGTGTGTGCTGCGGAGGCGCGGACCGCGCGGTGAGATCAAGATGCGCGCG 1068
DB 141 GAGAGCAGAGTGTGTGCGCGGTGAGCGGTCTGCGGAGGATGATCAGAGAGCAGCGCGAG 82
QY 1069 ATTCGGTCTGCGGTGCGAAGCGCAGCGCATCAGCGCAGCGAGCGCGCGCGCGCGCGCG 1128

```

```

DB 81 GTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22
QY 1129 GATCAGCAGATCATGCTCA 1148
DB 21 CAGCAGCAGCTATGCTCA 2

RESULT 7
AD096933/C
ID AD096933 standard; DNA; 1966 BP.
XX
AC AD096933;
XX
DT 23-SEP-2004 (first entry)
XX
DE CrtWcrty nucleotide sequence.
XX
KM carotenoid; transgenic plant; overlapping extension PCR amplification;
KM PCR; crtB; crtI; crtW; crtZ; astaxanthin; carotene;
KM grain crop; vegetable; crtWcrty; gene; de.
XX
OS Unidentified.
XX
PN CN1380415-A.
XX
PD 20-NOV-2002.
XX
PF 06-APR-2001; 2001CN-00105878.
XX
PR 06-APR-2001; 2001CN-00105878.
XX
PA (SHAN-) SHANGHAI YONGYE NONKO BIO ENG CO LTD.
XX
PI Yao Q, Peng R, Xiong A;
XX
DR MPI; 2003-230997/23.
XX
DR P-PSDB; AD096834.
XX
PT Synthesis of related gene for producing carotenoid in transgenic plant,
XX useful for improving quality of grain crops and vegetable.
XX
PS Example 1; Fig 1-2; 36pp; Chinese.
XX
CC The present invention relates to a related gene for producing carotenoid
CC in transgenic plant. The invention utilizes an overlapping extension PCR
CC amplification process and uses six genes of crtB, crtI, crtW, crtZ
CC and crtZ as templates to synthesize the related gene for synthesizing
CC astaxanthin with plant preference code and can make a synthetic gene for
CC expression in a plant to produce the carotene substances of astaxanthin.
CC The synthetic gene can be used to improve the quality of grain crops and
CC vegetables. The present sequence represents a specifically claimed
CC crtWcrty nucleotide sequence, which is used in the exemplification of the
XX present invention.
XX
SQ Sequence 1966 BP; 507 A; 721 C; 416 G; 322 T; 0 U; 0 Other;
XX
Query Match 34.4%; Score 395.4; DB 11; Length 1966;
Best Local Similarity 60.1%; Pred. No. 3,6e-65;
Matches 698; Conservative 0; Mismatches 451; Indels 12; Gaps 2;
QY 1 TCATGCTCTCTCTGAGCAGAGGGGGGCTTCGGGCAAGCAGCGACGGCTTCGACAGCGG 60
DB 1958 TTAATGGTTTCTCTTGAAGAGGTGCTTCTGAGAGCATCTGATTTGCGGTCCGAGTGG 1899
QY 61 AATGGGCGGGCGCTCGGTGACATGCGAAGCCGGTGGCCATGTCAAGGCGCCGCAATA 120
DB 1898 GATTTGGCTTCCCGGTGACATCTGAGTTGCTGTGAGAGAGAGTCTACGCGTA 1839
QY 121 GAAGCGCTGATCAGCGCGCTGCGGAGGCGGTGAACCGCTGACAGCGCATAGCAGCG 180
DB 1838 GAATCTTTCAGATGAGACCGGTGTGATCTCTTGTGAGAGAGGTGTATCTTCT 1779

```

OY	181	GTCCGGCGGGAGACCGCGGAAAACAGATCCGGTTACGACCGCGAAGAACGGATCGGCATC	240
Db	1778	GCTGGTTCGAGACCTCTGAAGAGCATTTCTGTTAGAGATCTGAAGAAATCTGTCTTTCT	1719
OY	241	CGCGGATCTGATATGCGCCCAAGCCGCGCACCGC-----GCCACGGGCGGACGGCGATCTGAC	294
Db	1718	TGCTCTGTGATATGCGATGTGTCTTGAATTTGACACTCTGATGTCGATCGGATCGCTGGAGACC	1659
OY	295	GTCCGCGCCCGCGATGCGATCCGCGACTCTGGCGGGCATATGGGACGCAATATCCGATAC	354
Db	1658	GGAGAGACCTCGAGACGACGTCTGCATCTTGCTGCTGCATATGGAGAGGAGTACCGCGTAC	1599
OY	355	GGGGTGAACAAGCCCTGCCCCCAGCCCAACCGGACACCGCCCCCTGCGGTGTGCGCGCA	414
Db	1598	TGGGTGAAGAAGCCTTGCTCTGAGACCGACCTGGACCTGACCTTGTGCTGTGATCTGCCCCA	1539
OY	415	GAAGCCTATGCGGTGATAGGCGCAGCGCATGGGACAGATATGCCCTTTGCGCCCGCATCTC	474
Db	1538	GAAACCTCTCTGCTGTGTGTGCGAAGTGCATTTGGAGAGATATGCTCTTTCTTTTGACTTC	1479
OY	475	CTGCCCGGTCCAGACCCCGCCCTGGCGGATAGTCAGCGACGCTTGCAGCGCGCATC	534
Db	1478	TGCACCGGTCCAGCCTTGATGTGATCTTGCGAGAGCTTGAGGATGTCGTGCGAGTGGCTGTG	1419
OY	535	GTCCAGATATCGCCCGCTGCTGTAGCGGGATATCTTCATTCAGAGATCGGGTGGACTGAA	594
Db	1418	GTCCAGGTGCGCGCGCTGAGAGTATCTGTGTCTTCGATGAGGAATTTGTGTTCTGGAGAA	1359
OY	595	GGGACGACGATATAGATGAAGCGGTAACCGTCCATCTGCGGAAACGGTCCGCTCATATGAT	654
Db	1358	TGGAGAGAGTATGATGAAGCCGTGACCGTCTCTTTGGGATGACGGTCCGCTCATATGATAC	1299
OY	655	CGGGCGCTCGACGCCATGGGAGGCGTGGTCTGATCTGACAGCCGCAAGAAATTTCTGGA	714
Db	1298	TGGTCTTGAGACACCGTGTGGCTGTGTGAGTTTCAATTTGACACCGAGAACTTGTTGAA	1239
OY	715	AACCAACGTCAGATGCGGGGATCTCGACGGACACACGGGCGTTCATCAACGACGACCTC	774
Db	1238	GCCGAGGATGAGGTCTCTGATATGGTATGAGGCTCTTGCCGCTCGAGAGAGAGCCGTCTTC	1179
OY	775	GATCCGCGAGCGCGTCCAGCGTGGCGCGGATATGTCACAGCTTCGCGACATGCGATAT	834
Db	1178	GATTTTAATTAACCGCAGAGAGAGGATTTGACCTTTGTGCTCGTAGAGTGCATATGTCCGAGTC	1119
OY	835	CCACCGCAGATCGACACC-----CTGCAGACAGCCCGATTCAGCGCGCCGCTCGATGCA	888
Db	1118	CCATCTGATTTCTGACCCGGAATCTGACGACATCGCTCTCGAGATGCTGACCGTGAAGGGA	1059
OY	889	GCCATATGCTGTCTGTCAGCGCGCGCGAAATGTCGGGAAAACGGACCTTCTGATCTCTGCA	948
Db	1058	GCCGTAGACCGGTTGCGAGCTCTTTGCGTGTCTTGAGGAAATCGAATTTCTTGATCTGAGCCA	999
OY	949	TTCCGCGCGACGAATGGCGCGACAGCGCGGCGACCATTCGGCGGAAATCCGTGTCTGTG	1008
Db	998	GTTTCTCTTCTGAATGTGCTTGAGTCTTTCAGAGCCAGTGTGGAGGAAGATCTGGATCTGTG	939
OY	1009	GCAGGACAGAGTGTCTGTGTCGAGAGGGACCGGACCGCGCTCGACATCAAGATCGCGAC	1068
Db	938	GCAAGAACAGAGTGTACCTTCGGATGAGACCTGTGCGTGTGTGAGAGAGAGACTCTGAG	879
OY	1069	ATCCGCTTCGCGGTGCGGAACGGCAAGCGCGATTCAGCGCACCGGACAGCCCCGCGCCGC	1128
Db	878	GTCCTGTCTTGTGCTCTGAGTGTGCGAGTGCAGATGAGACCGTTTGGAGACCTGCACCTGC	819
OY	1129	GATCAGCAGATATATGGCTCAT	1149
Db	818	GAGAGGAGACGTGTGGTCTAT	798

RESULT 8
ADQ96833/c

ID
ADQ96833 standard; DNA; 1944 BP.

XX

AC		ADQ96833;	
XX			
DT	23-SEP-2004	(first entry)	
XX			
DE	CrtWcrty nucleotide sequence.		
XX			
KM	carotenoid; transgenic plant; overlapping extension PCR amplification;		
XX	PCR; crtB; crtI; crtW; crtz; astaxanthin; carotene;		
KM	grain crop; vegetable; crtWcrty; gene; ds.		
XX			
OS	Unidentified.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1944	
FT		/tag= a	
XX		/product= "crtWcrty"	
PN	CN1380415-A.		
XX			
PD	20-NOV-2002.		
XX			
PF	06-APR-2001; 2001CN-00105878.		
XX			
PR	06-APR-2001; 2001CN-00105878.		
XX			
PA	(SHAN-) SHANGHAI YONGYE NONGKO BIO ENG CO LTD.		
XX			
PI	Yao Q, Peng R, Xiong A;		
XX			
DR	WPI: 2003-230997/23.		
XX	P-PSDB; ADQ96834.		
PT	Synthesis of related gene for producing carotenoid in transgenic plant,		
XX	useful for improving quality of grain crops and vegetable.		
PS	Claim 2; Page 1-2; 36pp; Chinese.		
XX			
CC	The present invention relates to a related gene for producing carotenoid		
CC	in transgenic plant. The invention utilizes an overlapping extension PCR		
CC	amplification process and uses six genes of crtB, crtI, crtW, crtz		
CC	and crtz as templates to synthesize the related gene for synthesizing		
CC	astaxanthin with plant preference code and can make a synthetic gene for		
CC	expression in a plant to produce the carotene substances of astaxanthin.		
CC	The synthetic gene can be used to improve the quality of grain crops and		
CC	vegetables. The present sequence represents a specifically claimed		
CC	crtWcrty nucleotide sequence, which is used in the exemplification of the		
CC	present invention.		
XX			
SQ	Sequence 1944 BP, 497 A; 716 C; 412 G; 319 T; 0 U; 0 Other;		
	Query Match	34.3%; Score 394; DB 11; Length 1944;	
	Best Local Similarity	60.1%; Pred. No. 6.7e-65;	
	Matches 696; Conservative 0; Mismatches 450; Indels 12; Gaps 2;		
Oy	4 TGCTTCCTCCTGCAGCAGGGGCGTTGGGCAAGGCGCACCGGCTTGCACAGCGAAT 63		
Dd	1944 TCGCTTTTCCTTAGAGAGTGCTCTTCTCGGAGGCGATCTGAATTCCGGTCCGAATGGGAT 1885		
Oy	64 GGGGGGGGGGTCCGGTGCAGATGCCAAAGCGGGTCCCAATTCAGGGCGCCGGCATGAA 123		
Dd	1884 TGGTGGCTTGGCGGAGCATTTTGAAGTTGGTGTCTGCGCAGGAGTGTTACTTGGGTAGA 1825		
Oy	124 GGCTGCATCAGCGGCTCGGCGAGGCGGTAGAACCGCTTGACGACGCGGATAGCAGCGT 183		
Dd	1824 TCTTTGCATGAGACCCTGTGGCATTTCTGTAAAGCTTTTGGAGGAGGGGTATCTTTGTC 17655		
Oy	184 GGGCGGCGAGCGCGGGAACAGCATCCGTTTACGACGCCGAGAAAGCGGTGCGATCCGC 243		
Dd	1764 TGGTGCAGGCGCTTGAAGAGCATTTGTGTAGGAGAGTCTGAGGAATCTGTCTTCTTGC 17050		
Oy	244 GCGATTCATGGCCAGCCGCGCACCC-----GCGACGGCGGAGCGGTCGTCAGGTC 297		
Dd	1704 TCTGTGAGTAGTGTAATCTGTGATTTGACCTTGAAGTGCCTGCGATGCTGTGATGACCGGA 1645		

OY	298	GGCGCCCGGAATGGGATCCGAGACTGGCGCGCATATGAGGAGACGAAATATCCGGTGA	CGGG	357
Db	1644	GAGACTCTCGACGACGCTCTGCGACTTTGGCTGCGATGAGAGGGGTACCGGATCTGG	1585	
OY	358	GTGGAACAGCCCTGCCCCAGCCCAACCGGACACCGCCCCCTGGCGGTGGTCCGCGACAGA	417	
Db	1584	GTGGAAAGAAAGCTGCTTGAGACCGACTGGGACTGTGACCTTCTGCTGGTCTGCCACAGA	1525	
OY	418	GCCTATGGCGTCATATGGGCGAGCGCATATGGGACAGATATGCCCTTTTGGCGCGCATTCCTG	477	
Db	1524	ACCTGCTCCTGCTGTGTGAGAGTGCATTTGGAGAGATGCGCTCTTCTCTTGACTTCTGCG	1465	
OY	478	CCCGGTCAGGCCCGCGCTGGGGGACATATGACAGGAGCGCTGGCGCAGGGCCATCGTC	537	
Db	1464	ACCGGTCAGGCTTTTGTCTTGGCTAGTCTTGGGATGTGCTGGAGTGGTGGTGGTGGT	1405	
OY	538	CAGATCGCGCGCGCTGCTGTATGCGCGATATCTTCGATCAGATATGCGGGTGGGACTGGAAGG	597	
Db	1404	GAGGTGCGCGCGCTGCGAGTATCTGTGTCTTCGATGAGAGATCTCGGTTCTGGAGAAATGG	1345	
OY	598	CAGCAGATATGATGAAGCGGTACCCGTCCATCTGCGGAAACGGTCCGCTCATATGATATGG	657	
Db	1344	GAGGAGGTAGATGAAGCCGTAGCCGCTCTCTTGGGTGACGGGTCCGTCAATGATACATGG	1285	
OY	658	GCGCTCGACGCCATATGGGGGGCGTCGGTCTCGATCTCGACGCCCAACGAATTTCTGGAAC	717	
Db	1284	TCTTGGGACACCGTGTGGTCTGTGCGTTTTCGATTTTCGACACCGAGAACTGTGGAAC	1223	
OY	718	CACGGTCAGGTGCGGGGATCTCGACCGGACCAACCGGCGTGCATCAGCAGGACGCTCGAT	777	
Db	1224	GAGGGTGAAGTGTCTGATATGGGTGTGGGCTCTGCGCGTCGAGGACGAGCGCTGCTCGAT	1165	
OY	778	CCGGAGCGCGTCCCTCAGCGCTGCGCCCGGTATCTGTTCAGCGCTGCGGACATATGCTATTCGA	837	
Db	1164	TCTATGTAACCGAGAGAGGGTTGACCTTGTGCGTCGAGAGTGGATGTGGAATCCCA	1105	
OY	838	CCGCGATCGACACC-----CTGCAGACGCCCGATCAGCGCGCGCGCTCGATGAGACC	891	
Db	1104	TCTGATTTCTGACCCGGATCTGACGACTGCGCTGTGCGAGTGTGACACCGTGAGGAGCC	1045	
OY	892	ATAGCTTGTCGTCAAGCGCGCGGAATGCTCGGGAAACCGGACCTTCGTATCCGTCAATTC	951	
Db	1044	GTAAGCGGTTGAGAGTCTTCTTTCGATCTTGGGAAATCGACTTCTTGATCTGGCCAGTT	985	
OY	952	GCCCGACGAAATGGGCGGACAGCGCGCGGACGATTCCTGGCGGAAAGATCCGTGTGTGGCA	1011	
Db	984	TGCTCTTCTTGTAGTGGCTTGTGAGTCTTGTGCAGGCAAGTGTGGGAGAGTGTGGGTGCTGGCA	925	
OY	1012	GGACCAAGTGTGCTGTGCTCGAGAGGGGCGGACCGCGCGTCGAGACATCAGATGCGCGCATTC	1071	
Db	924	GGACCAAGTGTGACCTTCTCGATGACCTGTGCGGTGTGCGAGAGAGAGAACTTGAGAGTTC	865	
OY	1072	CGGTTCGCGTTCGCGAAGCGGCAAGCGCGCATATGACGACCGGACAGACCCCGCGCGCAT	1131	
Db	864	TGGTCTTCTGCTCTCGAGTGCAGGTCGATGAGACGTTTTCGAGAACTTCGACCTTGCGAG	805	
OY	1132	CAGCAGATCATGCTCAT 1149		
Db	804	GAGGACGTGTGGGTCTAT 787		
RESULT 9				
ADC9013/c				
ID	ADC9013 standard; DNA; 1185 BP.			
XX	ADC9013;			
XX	AC			
XX	01-JAN-2004 (first entry)			
XX	Bradyrhizobium sp crty DNA.			
DE				
XX				
XX	gene cluster; canthaxanthin biosynthesis; carotenoid; astaxanthin;			

KW	poultry; fish feed; crty; de.
XX	
OS	Bradyrhizobium sp.
XX	
PN	US2003087337-A1.
XX	
PD	08-MAY-2003..
PF	
PE	11-JUN-2002; 2002US-00166037.
XX	
PR	12-JUN-2001; 2001US-0297272P.
XX	
PA	(GIRA/) GIRAUD E.
XX	
PA	(HANN/) HANNIBAL L.
XX	
P1	Giraud E, Hannibal L;
DR	
XX	
PT	Novel isolated gene cluster involved in canthaxanthin biosynthesis,
XX	
PT	having crty, crtI, crtB and crtW genes clustered in this order, in the
XX	
PT	same orientation, and crtE preceding 4 genes in the opposite direction.
PS	
XX	
PS	Claim 6; SEQ ID NO 2; 14pp; English.
CC	
CC	The invention relates to a novel isolated gene cluster involved in
CC	canthaxanthin biosynthesis comprising a polynucleotide in which crty,
CC	crtI, crtB and crtW genes are clustered in this order and in the same
CC	orientation and the crtE gene is oriented in the opposite direction. The
CC	polynucleotide of the invention may be useful for producing at least one
CC	carotenoid such as astaxanthin or canthaxanthin which may be used within
CC	poultry or fish feed. The current sequence is that of the Bradyrhizobium
CC	sp. crty DNA of the invention.
CC	
SQ	
SQ	Sequence 1185 BP; 169 A; 414 C; 390 G; 212 T; 0 U; 0 Other;
	Query Match 20.6%; Score 236.4; DB 10; Length 1185;
	Best Local Similarity 54.0%; Pred. No. 2.5e-35;
	Matches 589; Conservative 0; Mismatches 471; Indels 30; Gaps 4
OY	
Db	78 TGACGATGCCGAAGCGGGTCCGCGCAATGCAGGCGCGCGCATAGAAGCGCTGCATCAGCG 137
Db	1102 TGAAGACGCGGAGCTTGTCCTCTGCGGCTGGAATCCGGGCGGCGGTGAAGCGCTGCATCAGG 104
OY	138 GCTTCGGGAGCGGTTAGAACCGCTGCAGACGCGCATAGCAAGCTCCGGGCGGCGACCGCG 197
Db	1042 CCTATGAGACCGTAATAATGTGTCAGAGATTCCGTGACGCTCGAAGGGCTTCGGGGCCT 983
OY	198 GGAACAGCATCCGTTTACAGACCGCGCAAGAGCGGTCCGCATCCGCGCGATGATGAGCCC 257
Db	982 TGAACACGATCCGGTTACGAAAGGATGTAGCCGCGCGCGCCAGATCTGCCGCCGT 923
OY	258 AGCGGCGCACCGCGCGACCGGCGGACCGGGTGTCTAGGTCG-----CGCG 302
Db	922 AGCCATGATGTCTTCGCGCGCGTCAAGGCTCAAGAGCGCGCGCCGTTGCGCAAGCCGG 863
OY	303 CCGGATATGACATCCCGGACCTGCAGCGCATAGGGCAGCGAATATCCGGTGAACGGGTGGA 362
Db	862 CCGTCAGCAAGATCGGCGACCGGACCGCGTCCGGACGGAATAGCCGATCTCGCATGGA 803
OY	363 ACAGCCCTGCCCCAGCGCCAAACCGGACAACCGCCCCCTGCGGTGTGCGCGCAAGACTTA 422
Db	802 CGAAGACCGCGCGAGACCACCGCGCGCGGGCTGT---CGGGCTTGTGACCAAGCCCG 746
OY	423 TTGGGTATAGGGCCAGCGGCATATGGCGAGAGATGCCCTTTTCCGCGCGCATCTTCGCCGG 482
Db	745 ACGGATGCGCGCGCAAGATGACGGGACCAACCTTGTCTCCGCGCGAATGATCTCGGCGCA 686
OY	483 TTCAGCCCGCGCTGAGCG---GCATATGTCAGACGACGCTTGCAGCGCGCGCATGTCCA 539
Db	685 TCTGCCAGACCTTTGGCAGACCGCGTATGCGCGGATGCGCTTATGACAGACCTGTCTGGGGA 626
OY	540 GATGCGCGCGCTGCTAGCGGTATCTTCATTCAGATCGAGATCGGGTGGACTGAAGGCA 599

Db 625 ACTCCGCGCTGCTATAGTAGTGCTCTCATCAACGCGCTGAGGCTGACCGCA 566
QY 600 GCAGATAGATGAAGCGGTACCCGTCATCTGGCGAAACGGTCCGTCATGATCGGCG 659
Db 555 GCGTGTAGACGAAGCGATAGCCGCTCTGCGCGACCGTCCGCTTCATCAAGTGGCA 506
QY 660 GCTCAGCGCATGGGGGGCGGTGCTGATCTGACGCGCCAGATTTCTGAAACCA 719
Db 505 CATTCAGAGCGGTGCGCGCGCGCGACCGCATCTCGAGCCCGAGAACTTGAACCGA 446
QY 720 CGGTACAGTGGCGGGTCTGACGCGACCGCGCTTCATCAACGAGAGCTTGATCC 779
Db 445 GAGCGAGCCCGCGACCGCGCGCGCGCGCATGATCAACGAGGCGCGCGCAAG 386
QY 780 GCGAGCGCTGCTGACG-----GTGCGCGCGTATCTGTCAGCGCTGCGACATGCG 830
Db 385 TCCGCGCGCTCTCCAGCAGACATGCTCTGCGGTGCGAGATGCGCGCTGCGCGCA 326
QY 831 TATTCACCGCAGATCGACACCTTGACGACGCCCATTCAGCGCGCGCTTGATCGAGC 890
Db 325 GGATCCTCTCGGAAAGCGCTGACCTCGGACGCGCGCGCTGACGTCATGAGGA 266
QY 891 CATGACCTGTGCTGACGCGCGCGCGATGCTGCGGAAACCGCATCTCTGATCCGTCATT 950
Db 265 GGTAGCGCGTTCAGACAGCGATGCGCATGTTTCGCGCAAGCGCATTCATTAACCGCGCAGC 206
QY 951 CGCGCGCAGATGAGCGCGACGAGCGCGCGCGCATTCGCGGAAAGATCGGTGCTGCGC 1010
Db 205 GATGACCGACGAGCGCTCGGACCGCGCTGCTGTCGACGAGATGCTGTCGCAAGA 146
QY 1011 AGGACGAGGTGTGCTGCTCGAGGCGCGCGCGCGCTGACGATCAAGTGCAGCAT 1070
Db 145 AGCTCAGGTGTGATGTCGCGCGCATGCTGCGCGAGCTTCGATGATGACGACGCGCAAT 86
QY 1071 CCGGCTGCGGTGCGGACGCGACGCGCATGACGCGACCGCGCGCGCGCGCGCA 1130
Db 85 CCGGCGCTGCTGCTGCTGACCGCGAGGCGATGAGCGCGCGCAAGCGCGCGATGA 26
QY 1131 TCAGCAGATC 1140
Db 25 CGATGACGTC 16

RESULT 10
ADM98599/c
ID ADM98599 strand: DNA; 12753 BP.
AC ADM98599;
XX
XX
XX 01-JUL-2004 (first entry)
DE Geranylgeranyl pyrophosphate synthase DNA #19.
XX
XX Geranylgeranyl pyrophosphate synthase; gene; ds; diterpene;
KM diterpene precursor; diterpene synthase; defence toxin;
KM volatile defensive signal; pollinator attractant; photoprotectant.
XX
OS Pantoea agglomerans.
XX
XX US2004072323-A1.
XX
XX 15-APR-2004.
XX
XX 07-JAN-2002; 2002US-00041018.
XX
XX 05-JAN-2001; 2001US-0259880P.
XX
XX (MATS/) MATSUDA S P T.
PA (HART/) HART E A.
XX
XX Matsuda SPT, Hart EA;
XX

DR WPI; 2004-373921/35.
XX
XX New unicellular organisms comprising exogenous nucleic acids encoding a
PT geranylgeranyl pyrophosphate and a diterpene synthase, useful for
PT producing diterpenes and diterpene precursors.
XX
XX Claim 2; SEQ ID NO 19; 38bp; English.
PS
XX
XX The invention relates to a unicellular organism for producing a diterpene
CC or diterpene precursor comprising an exogenous nucleic acid sequence
CC encoding a geranylgeranyl pyrophosphate synthase under the control of a
CC promoter operable in the organism, and an exogenous nucleic acid sequence
CC encoding a diterpene synthase, under the control of a promoter operable in
CC the organism. The invention also relates to methods of producing a
CC diterpene or diterpene precursor and a method of isolating a diterpene
CC synthase comprising growing several cells in the presence of a
CC polycyclic aromatic resin to make a cell/resin mixture, where at least one of
CC the cells further comprises at least one isolated and purified nucleic
CC acid sequence of a yeast expression library, and the expression of the
CC nucleic acid sequence is regulated by an inducible promoter under
CC conditions where the expression is induced, filtering the cell/resin
CC mixture, extracting the cell/resin mixture with alcohol to produce an
CC organic eluent and analysing the organic eluent by a screening method
CC including chromatography and/or spectroscopy, to identify the nucleic
CC acid sequence encoding the diterpene synthase. The unicellular
CC microorganism is useful as a diterpene or diterpene precursor producing
CC system. Diterpenes, in plants, serve as defence toxins, volatile
CC defensive signals, pollinator attractants and photoprotectants. This
CC sequence represents DNA encoding a geranylgeranyl pyrophosphate synthase
CC polypeptide of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
SQ Sequence 12753 BP; 2489 A; 3880 C; 3911 G; 2473 T; 0 U; 0 Other;
XX
XX
XX Query Match 18.6%; Score 213.4; DB 12; Length 12753;
XX Best Local Similarity 52.1%; Pred. No. 51e-31;
XX Matches 581; Conservative 0; Mismatches 516; Indels 18; Gaps 4;
QY 41 GCACGCGCTGCGACGCGAATGAGCGCGGTCCGCTGACGATCGAAGCGGTGCGCC 100
Db 7916 CGCCAGGCTTGCGCCAGCGGAACCGGTGCTGCCCTGCAAAATCGGGGCTTATCAAG 7857
QY 101 AATGTCAGCGCGCGCATGAAAGCTTCGATCAGCGGCTGCGCGCGGTGAAAGCC 160
Db 7856 AGAGAGACCGACCGCGCTTAAAGCGCTTACGCTGCGCGACCGCATTAAGCGCG 7797
QY 161 TGCACAGCGCATACCGACCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 220
Db 7796 TGCATCACCGCGACCGGTCTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 7737
QY 221 CGCAGGAAGC---GGTCGCGATCCGCGCATGAGCGCGCGCGCGCGCGCGCGCG 277
Db 7736 CGGAAGATCCCTGCTGCGCGCATGAGCGCGGTTCGGAACCTCGCGGTAGATGAGC 7677
QY 278 GCGACGCGGTGTCAGGTGCGCGCGCGCGCATGAGCGCGCGCGCGCGCGCGCG 337
Db 7676 GGAACGCTGCCAGCGCGCGCGGTGCGCGCATGCGGTGCGCGCGCGCGCGCGCG 7617
QY 338 AGCGAATATCCGCTGACCGGTGGAACAGCCCTGCGCGCGCGCGCGCGCGCGCG 397
Db 7616 AGCGAATGACGATGATGAGGTGGAATAGCCAGCGCGCGCATTCGCGAGCGCGCGCG 7557
QY 398 TCCGCTGATGCGCGCGCGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 457
Db 7556 GCGCATGCGC---CCACAGAGCTGAGTGTACCCGCGAGGTATGCGCAGAGCGCG 7500
QY 458 CTT---TGGCGCGCATCTCTGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 514
Db 7499 GTCTCCTGCGGTAAAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTAAC 7440
QY 515 GCTTGGCGCGCGCGCGCATGTCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 574


```
QY 579 TCGGGGTGGAGCTGAGGACAGATAGATGAGGATACCGCTCATCTGCGGAAACGG 638
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
PS GTGGGTTGGGAAAGCGGCGAGGATATAAACAAGGAGTAGCCATTTGCTGATGACCG 527
QY 639 TGGCGTCCATGATATCGGGGCGCTGACCGCATGGGGGGCGTCCGCTTCGATCTGACGC 698
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 526 TCGCATTCATGATATTCGGTGAAGATMAACATGGCGCGCGCTCAGTTGCCACTCTGAC 467
QY 699 CCACGAATTTCTGAAACCCACGGGTGCGGGGCTTCGACGGGACCAAGGGGCTGCA 758
DB 466 CCAATTAATCTTGGAATCTTACGCGTAGTGCAAGATCAGGCGGTMAACCCCTTCGTGCA 407
QY 759 TCACGAGGACAGCTTCGATCCGAGACCGCTCCGTGACGCTCGCGCGATTCGTCAAGC 818
DB 406 TCAGTGTACTGGCATGATATATCCGCGCATCCGCTTACGTGACCCGATTCAGATGAACGG 347
QY 819 TCGGCATATGCGTATTCACCGGATTCACACCTTCGACGACCGCGATCAGCGG---- 874
DB 346 CTGAAACCGCGGTATGACAGCCATMAATGTTGTCAAACCTGTTGCCGAGTATCCCGGCGA 287
QY 875 -----CCGCTCGATCGAGCCATAGCCCTGTGCTCAGCGGCGGCGAATGTCGGAAACG 929
DB 286 AATGCCGGAGAGTCAAGCATAGTAGTACCACTGTTCAATGCGGACGCGCTTGGGGAAAC 227
QY 930 CGACCTCTGATCCGTTCATTCGCCCGCAGCAAGATGGGCGCAGCGCGCCATTCGG 989
DB 226 GAACCTGTAGTTCGGGCGAGTGTGACCAAGCGGCGCTATCCAGCATGCTGATTCGA 167
QY 990 GCGAAGATTCGCTGTGTGACGAGACAGAGTGTGCTGTGTCGAGGGGCGGACCGCGCT 1049
DB 166 GGGTAAATCTCTTCTGTGAAAGACAGAGTATGTTCCCTCCGCTCAGAACCGCGCT 107
QY 1050 CGAGCATCAGATGCGCGCATCCGCTCTGCTGCGGAAACGGACCGGATCAGGCGAC 1109
DB 106 CAATTAAGCAAGATCCGATATTCGATGCTGTGCTGAAGCCGGAGCGGATTAAGCCAT 47
QY 1110 CGGACAGCCCGCGCGCGCGGATCAGCATCATGG 1144
DB 46 TAGCCGAGCCGAGCCGACCAAGATGATCATAG 12

RESULT 12
ABT14191/c
ID ABT14191 standard; DNA; 1149 BP.
AC ABT14191;
XX
XX 20-FEB-2003 (first entry)
XX
XX Pancoea stewartii lycopene Beta-cyclase gene.
DE Pancoea stewartii.
XX
XX Gene; ds; carotenoid; crt.
XX
XX Pancoea stewartii.
XX
XX WO200279395-A2.
XX
XX 10-OCT-2002.
XX
XX 25-JAN-2002; 2002WO-US002124.
XX
XX 26-JAN-2001; 2001US-0264329P.
XX
XX 04-MAY-2001; 2001US-0288984P.
XX
XX (CRGI ) CARGILL INC.
XX
XX De Souza ML, Kollmann SR, May CA, Schroeder WA;
XX
XX MPI; 2003-075455/07.
XX
XX P-PSDB; AAO16019.
XX
XX Novel isolated nucleic acid useful e.g. to engineer host cells with the
PT ability to produce particular carotenoids and polypeptides useful in cell
```

```
PT -free systems to make particular carotenoids.
XX
XX Claim 81; Page 60; 74pp; English.
PS
XX
XX The invention comprises the amino acid and coding sequence of a number of
CC carotenoid (crt)-related proteins. The crt-related DNA and protein
CC sequences of the invention are useful for engineering cells which are
CC able to produce carotenoids. The present DNA sequence represents a crt-
CC related gene of the invention
XX
SQ Sequence 1149 BP; 228 A; 325 C; 315 G; 281 T; 0 U; 0 Other;

Query Match 18.3%; Score 209.8; DB 8; Length 1149;
Best Local Similarity 51.8%; Pred. No. 2.5e-30;
Matches 578; Conservative 0; Mismatches 522; Indels 15; Gaps 4;

QY 42 GCACGCGCTGGAGACAGCGGAATGGGCGGGCGCTCCGGGTGACGATGGAGACCGGTGCGCA 101
DB 1123 GCATGCGCGGAAACGCGAAACGGGCGGCTTGGCGCTCAGAAATGCTAGCGGATGCTCA 1064
QY 102 ATGTACGCGCGCGGATAGAAAGCGCTTCATCAGCGGCTGCGGCGAGCGGTAGAACCGCT 161
DB 1063 CGGTAAATTTTCCGCAATAAAGCGGCAATCAATCCTCGGGTAAAGCCATTAAGACGCT 1004
QY 162 GCAGCAGCGCATAGCAGCGGTGCGGCGGCGAGCGCGGAAACAGCATCCGCTTCAGACGC 221
DB 1003 GCATCAGACGCGCAGCGTACTCGGCGGCTCCGCTTAACAAACATGCGATTCAGCATGC 944
QY 222 GCAGGAACGGGTCCGGATCCGCGGATCGATGAGGCCAG---CCGCGCACCGGCGGACGGG 278
DB 943 GGAAGAACCCCTGTTGCTGTCACAGTCTGTGGGCAAAATGAGCAATCTGTGTGTAACAG 884
QY 279 CGAGCGCGGTGTCAGTTCGCGCGCGCGGATGAGCATCCGCACTGCGGCGCATAGGCGCA 338
DB 883 AGGAAGAGTAAACATATCAGCGCGGCTGAGACATATGCGGCGGCGGCGGCGGCGGTA 824
QY 339 GCGAATATCCGGTGACGCGGGTGAACACGCTTCGCCCAAGCCCAACCGGACCGCCCT 398
DB 823 GGAATATCCGCTGTGTGCGAATGAACACGCGGCGGCTTAATCCGTACAGGCTTGGGTT 764
QY 399 GCGCGTGTTCGCGCCAGAAAGCTATGCGGTATAGGCGGCGGCGGATGAGGCGCC 458
DB 763 GC-TGTTGCCAAATGACGATTAATGCGCGCTTAAAGTAATGAGCAATGACCTGTTTC- 706
QY 459 TTTCCGCGCGCATCTCTGCGCGGTTCAGCGCGCGCTGCGGCGCATATGTCAGCGACGCT 518
DB 705 -TTCCCGACAGCACTCTGTAAACGCGCAACCTGTGCGGCGGCAATATGCGGAATGTTCT 647
QY 519 GCGCCAGCGCGCATGCTCAGATGCGCGCGGTGCTGTAGCGGCTATCTTCGATCAGGA 578
DB 646 GACGCGCGCGTTCGCGCTGAAGATTAGCTTGTCAATGTAGTGTGCTTCGATCAGCA 587
QY 579 TCGGGGTGGAGCTGAGGACAGATAGATGAGGAGGATCCGCTCATCTGCGGAAACGG 638
DB 586 GTGGGTTGGGAAAGCGGCGAGGATTAACAAGCGGTAAGCATTTTGTGCTGATGACCG 527
QY 639 TGGCGTCCATGATATCGGGGCGCTTCAGCGCATAGGGGGCGTCCGCTTCGATCTGACGC 698
DB 526 TCGCATTCATGATATTCGGTGAAGATMAACATGGCGCGCGCTCAGTTGCCACTCTGAC 467
QY 699 CCACGAATTTCTGAAACCCACGGGTGCGGGGCTTCGACGGGACCAAGGGGCTGCA 758
DB 466 CGATTAATGCGCTGGAATCTTACGCGTATGCAAGATCAGGCGGTGTAACCCGCTTCGTGCA 407
QY 759 TCACGAGGACAGCTTCGATCCGAGACCGCTCCGTGAGGCTGCGCGGATTCGTCAAGC 818
DB 406 TCAGTGTACTGGCATGATATATCCGCGCATCCGCTTACGTGACCCGATTCAGATGAACGG 347
QY 819 TCGGCATATGCGTATTCACCGGATTCACACCTTCGACGACCGCGATCAGCGG---- 874
DB 346 CTGAAACCGCGGTATGACAGCCATMAATGTTGTCAAACCTGTTGCCGAGTATCCCGGCGA 287
QY 875 -----CCGCTCGATCGAGCCATAGCCCTGTGCTCAGCGGCGGCGAATGTCGGAAACG 929
```



```

Db      286 AATCGCGGAGGTACGACGAGTAGTACCACTGTTTCATGATCGGACCGCGTGGGGAAAC 227
QY      930 CGACCTTCATTCCTGTCATTCGCCCGGACGAGATGGCGGACGCGCCATTCGG 989
Db      226 GAACCTGTGAGTGGGCGAGTAGGACCAACGCGCGCTATCCGAGTGTGATTC 167
QY      990 GCGAAGATCCGATGTCGTCGACGACGAGTGTGTCGATCCGAGGGGCGGACCGCGGT 1049
Db      166 GCGTTAACTCTTGTGTGAAGACCAAGTATGTTCTCCCGCTCAGACCCGCT 107
QY      1050 CGAGCATCAGATCGCGATCGGATCGGTCGCGTCCGAAACGCGACGCGCATCGCGAC 1109
Db      106 CAATAGCAAGATCCGATATCCGATGTGTTGTGTAAGCCGAGGCGCATAGGCAAT 47
QY      1110 CGGACGCCCCCGCGCGGATCAGCATCATCG 1144
Db      46 TAGCCAGACCGGACCGACCGAATGATCATATG 12

RESULT 13
ACC44761/c
ID      ACC44761 standard; DNA; 1149 BP.
AC      ACC44761;
XX
XX      03-JUN-2003 (first entry)
XX
XX      Pantoea stewartii lycopene cyclase encoding DNA SEQ ID NO:5.
DE
XX      Pantoea stewartii; carotenoid biosynthetic enzyme; crtB; crtX; crtY;
KM      crtI; crtB; crtZ; lycopene cyclase; enzyme; phytoene; carotenoid; gene;
KM      ds.
OS      Pantoea stewartii.
XX
XX      Key      Location/Qualifiers
FH      CDS      1..1149
FT      /tag= a
FT      /product= "lycopene cyclase (crtY)"
FT      /note= "carotenoid biosynthetic enzyme"
XX
XX      MO2003016503-A2.
XX
XX      27-FEB-2003.
XX
XX      15-AUG-2002; 2002WO-US026647.
XX
XX      15-AUG-2001; 2001US-0312646P.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
XX      Brzostowicz PC, Cheng Q, Picataggio SK, Rouviere PE;
PI      WPI; 2003-268323/26.
XX
XX      P-PSDB; ABP96687.
XX
XX      Novel nucleic acid molecule isolated from Pantoea stewartii encoding a
PT      carotenoid biosynthetic enzyme, useful for regulating carotenoid
PT      biosynthesis in an organism.
XX
XX      Claim 2; Page 60-61; 68pp; English.
XX
XX      The present invention describes Pantoea stewartii carotenoid biosynthetic
CC      enzymes (1). More specifically described are the geranylgeranyl
CC      pyrophosphate synthase (crtB), zeaxanthin glucosyl transferase (crtX),
CC      lycopene cyclase (crtY), phytoene desaturase (crtI), phytoene synthase
CC      (crtB) and beta-carotene hydroxylase (crtZ) enzymes (see ABP96686 to
CC      ABP96690) encoded by ACC44759 to ACC44764. (1) can be used for regulating
CC      carotenoid biosynthesis in an organism. By over-expressing (1) in an
CC      organism, such that the carotenoid biosynthesis is altered in the
CC      organism. (1) and the genes encoding (1) are useful for converting
CC      phytoene to the carotenoids, for creating recombinant organisms that have

```

```

CC the ability to produce various carotenoid compounds, and also for
CC enhancing or manipulating carotenoid compounds. (1) can also be used for
CC producing gene products having enhanced or altered activity
XX
XX      Sequence 1149 BP; 228 A; 325 C; 315 G; 281 T; 0 U; 0 Other;
SQ
Query Match      18.3%; Score 209.8; DB 8; Length 1149;
Best Local Similarity 51.8%; Pred. No. 2.5e-30;
Matches 578; Conservative 0; Mismatches 522; Indels 15; Gaps 4;

QY      42 GCACGCGCTTCGACACGCGAATGGCGGGCGTCCGTCGATCGACATCGGACCGGTGGCCA 101
Db      1123 GCAATGCCCGCAAAACGGGAGCGGGCGCTTCCGCTCAGAAATGCGTACCGCATCGTCA 1064
QY      102 ATGTACGGGCGCCCGCATGAAGCGGTGATCAGGGGCTGGGCGGAGGGGTGAACCGGT 161
Db      1063 CGGTGAGTTTTCCTCGATAAAGCGGGCATATAATCTCGGGTAAAGCATGAAGACGT 1004
QY      162 GCAGCAGCGATAGCAGCGGTGCGGCGGACCGCGAACAAGCATCCGGTTCAGCAGCC 221
Db      1003 GCATCAGACGCGACGCTGACTCGCCGCTCGGCTAAACAAACATGCGATTCAAGATCG 944
QY      222 GCAGGAAGCGGTGCGGATCGCGCATCGATGCGCCAG---CCGCGCACCGGCGACGGG 278
Db      943 GGAATAAACCCCTGTTGCTGCTCAACGTTGCTGGGCAAAAGTACCAATCGTCTGGTAA 884
QY      279 CGAGCGCGGTGCTGATGTCGCGCGCGCGATGCGATCCGCGACCTGCGGCGATAGGGA 338
Db      883 AGGAAGAGGTAAACATCATCAGCGCGCTGAGACGATCGCGCGCACCGGACGGAGCGTA 824
QY      339 GCGAATATCCGGTGAACGGGGGTGAACAGCGCTCCGCCCAACCGGACCGGCCCCCT 398
Db      823 GGGAGTAGCGGCTTGTGCGATGAACAGCGCCGCGGTATCCGCTACAGGCTTGCGTT 764
QY      399 GCGCGTGTGCGCGCAAGACCTATGCGCTCATGCGGCGCAGCGCGATGGCGAGATCCCC 458
Db      763 GC-TGTTGCCAAATGACAGATTATCGCCGTTAAGTATGGGCAATGACCTGTTTC- 706
QY      459 TTTTCGCGCGCATCTCCGCCCGGTCGACCGCCGCTGGCGCGCATATGTCACAGCGCT 518
Db      705 -TTCCCGACGACACGTGTGTAACGGCCCAACCTGTGCGCGACCATTAATGCGCAATGTTCT 647
QY      519 GCGCAGCGCGCGCATTCGTCAGATCGCGCGCGCTGCTGATGGCGGTATCCTGATCAGGA 578
Db      646 GACGCGCCCGCTTCGCGCTGAAGATTAGCTTGTCAATGATGATGTGTCTTGATCAGCA 587
QY      579 TCGCGGTGGACTGAAGGCGCAGACAGATGATGAAGACCGGTATCCATCTGCGGAACGG 638
Db      586 GTGCGGTTGCGAAGGCGCAGGATTAACAAACCGGATGACATTTTGCTGATCGACCG 527
QY      639 TCGCGTTCATGATCATCGGCGCGCTGACCGCATGGGGGCGCTGCGTCTGATCTGACCG 698
Db      526 TCGCATTCATGATTAATCGGTGACGATTAACCATGCGGCGCGCTCATGTTGCCACTCGAC 467
QY      699 CCACGAATTTCTGAAGAACCCACAGGTCAAGGTGGGGGTCTCGACGCGACGAGGGGCTGA 758
Db      466 CCAATAAATGCTGGAATCTTACGCGTAGTGAAGATCAGGCGTGAACCCGCTGCTGA 407
QY      759 TCACGAGGACGCTCGATCCGCGAGCGCTCGTCCGTCGCGCGGTATGTCACAGG 818
Db      406 TCACGTGATCGCATGAATTAATCGGCGCATCGGTATCGACGCGATTCAGATGAACGG 347
QY      819 TCGCATATGCGTATTCACCGCAGATGACACCTCGACAGAGCCGATCAGCGG---- 874
Db      346 CTGAACACCGCGGTATGACGCGCAATATGTTGTCAAACTGTTGCCGAGTATCCCGGCGCA 287
QY      875 -----CCGCGCTCGATCGACGATGCGCTGTCGAGGCGGCGGAATGGTGGGAAACG 929
Db      286 AATGCCGGAGGTACGCGAGTAGTAGTACCACTGTTTCATGATCGGACCGCGTGGGGAAAC 227
QY      930 CGACCTTCATTCCTGTCATTCGCCCGGACGAGATGGCGGACGCGCCATTCGG 989
Db      226 GAACCTGTGAGTGGGCGAGTAGGACCAACGCGCGCTATCCGAGTGTGATTC 167

```

Oy	990	GC	GAAGATCCGTCGCGGACAGACCAAGTGTGCTGGTCCGAGGAGCGCGACCGCGGT	104
Db	166	GC	GTTAATCTCTTCGTGAAAGACCAAGATGTCTTCCTCCGCTCAGAACCCGCTT	107
Oy	1050	CG	AGCATCAGCATGCGCGCATTCGGTCTCGCGATCGGAAACGACGCGATCAGCGCAT	1109
Db	106	CA	ATTAAGCAAGATCCGATATCCGATGTCGTGTGGTAAAGCCGAGCGCGATTAAGGCAT	47
Oy	1110	CG	ACAGCCCGCGCGCGCATCAGCAATCATGG	1144
Db	46	TAG	CCAGACCGGACCGACCGACAGAAATGATCATAG	12
RESULT 14				
ADQ14628/c				
ID	ADQ14628	standard; DNA; 1149 BP.		
XX	AC	ADQ14628;		
XX	DT	23-SEP-2004 (first entry)		
DE	XX	Pantoea stewartii lycopene cyclase encoding DNA SEQ ID NO:5.		
XX	XX	carotenoid; carotenoid overproducing microorganism;		
KW	XX	functional isoprenoid enzymatic biosynthetic pathway; Pantoea stewartii;		
KW	XX	crty; gene; lycopene cyclase; enzyme; ds.		
XX	XX	Pantoea stewartii.		
OS	XX	Key	Location/Qualifiers	
FH	XX	CDS	1..1149	
FT	XX		/*tag= a	
FT	XX		/product= "lycopene cyclase"	
XX	XX	WO2004056974-A2.		
PN	XX	08-JUL-2004.		
PD	XX	19-DEC-2003; 2003WO-US041811.		
PF	XX	19-DEC-2002; 2002US-0435612P.		
PR	XX	(DUPLO) DU PONT DE NEMOURS & CO E I.		
XX	XX	Cheng Q, Rouviere PE, Tao L;		
PI	XX	WPI; 2004-525436/50.		
DR	XX	P-PSDB; ADQ14629.		
DR	XX	Novel carotenoid overproducing microorganism comprising genes encoding		
PT	XX	functional isoprenoid enzymatic biosynthetic pathway comprising disrupted		
PT	XX	genes such as dead, mreC, and yfhE, useful for producing carotenoid.		
XX	XX	Example 2; SEQ ID NO 5; 85bp; English.		
XX	XX	The present invention describes a carotenoid overproducing microorganism		
CC	CC	comprising the genes encoding a functional isoprenoid enzymatic		
CC	CC	biosynthetic pathway comprising a disrupted gene chosen from dead, mreC,		
CC	CC	and yfhE. Also described is a carotenoid overproducing Escherichia coli		
CC	CC	microorganism (II) comprising an upper isoprenoid enzymatic biosynthetic		
CC	CC	pathway comprising the genes dsx, dxt ygbP, ychB, ygbB, lytB, idi, ispA		
CC	CC	and isgB, a lower isoprenoid enzymatic biosynthetic pathway comprising		
CC	CC	the genes crtE, crtB, crtI, and crty, mutations chosen from a mutation in		
CC	CC	the tms gene of the 3159 nucleotide sequence of SEQ ID NO:35 (ADQ14658),		
CC	CC	mutation in the rpsA gene of the 2904 nucleotide sequence of SEQ ID NO:37		
CC	CC	(ADQ14660), mutation in the rpoC gene of the 5454 nucleotide sequence of		
CC	CC	SEQ ID NO:38 (ADQ14661), mutation in the yjeK gene of the 1845 nucleotide		
CC	CC	sequence of SEQ ID NO:39 (ADQ14662), or mutation in the rhl gene of the		
CC	CC	2676 nucleotide sequence of SEQ ID NO:41 (ADQ14664), where the genes of		
CC	CC	the lower isoprenoid enzymatic biosynthetic pathway reside on an		
CC	CC	autonomously replicating plasmid comprising a replicon chosen from p15A		
CC	CC	and pM1. (I) chosen from bacteria, yeasts and filamentous fungi such as		

	CC	Aspergillus, Trichoderma, Saccharomyces, Pichia, Candida, Hansenula,
	CC	Salmonella, Bacillus, Acinetobacter, Zymomonas, Agrobacterium,
	CC	Erythrobacter, Chlorobium, Chromatium, Flavobacterium, Cytophaga,
	CC	Rhodospirillum rubrum, Rhodococcus, Streptomyces, Brevibacterium, Corynebacteria,
	CC	Mycobacterium, Deinococcus, Paracoccus, Escherichia, Erwinia, Pantoea,
	CC	Pseudomonas, Spingomonas, Methylobacter, Methylococcus,
	CC	Methylosinus, Methylothermobacter, Methylophilus, Alcaligenes,
	CC	Synechocystis, Synechococcus, Anabaena, Thiohalobium, Stephylococcus,
	CC	Methanobacterium, Klebsiella, and Myxococcus, and (II) are useful for
	CC	producing carotenoids which involves contacting (I) or (II) with a
	CC	fermentable carbon substrate, growing the carotenoid overproducing
	CC	microorganism for a sufficient time to produce carotenoid, and optionally
	CC	recovering the carotenoid from the carotenoid overproducing
	CC	microorganism. The present sequence represents the Pantoea stewartii cty
	CC	gene which encodes lycopene cyclase, which is used in an example from the
	CC	present invention.
	CC	
	XX	
	XX	
	XX	
	Sequence	1149 BP; 228 A; 325 C; 315 G; 281 T; 0 U; 0 Other;
	Query Match	18.3%; Score 209.8; DB 12; Length 1149;
	Best Local Similarity	51.8%; Pred. No. 2.5e-30;
	Matches	578; Conservative 0; Mismatches 522; Indels 15; Gaps 4
OY	42	GACAGCGCCTGGGACACGACGGAAATGGGCGGGCGGCCTGGGTACGATGCGGAAGCGGTGCGCA 101
Dd	1123	GCAATGCCCGCAAACGGGAACGGGCGGCTTGCCTCATGAATGCGTAGCCGATCGGTCA 106
OY	102	ATGTCAGGCGCCCGGCATAGAAGCGCTTGATCAGCGGCTGCGGACGAGCGGTAGAACCCT 161
Dd	1063	CGGTGAGTTTCCCGCATMAAAGCGGCGCATCAATCTCGGGTAAGCCATAGAACAAGCT 100
OY	162	GCAGCAGCGCATAGCGACGCGTCGGCGGGGACGCGCGGAACAGCATTCGGGTTCAGACGC 221
Dd	1003	GCATCACACGCGCAGCGTACTCGCGCGGTCCGGCTMAAACAACATYGGAATTCAACATGC 944
OY	222	GCAGGAACCGGTCGGGATCCGCGGATGATGAGCCAG---CCGCGACCGGCGCACGCG 278
Dd	943	GGAAAAACCCCTGTGTCGTCGCAACGTTTCTGGGCAAAETGAGCAATCTCTGTGTGAACG 884
OY	279	CGAGCGCGGTGTCAGGTTCGCGCGCGCGGATGAGCATTCGCGACCTGCGCGCATAGGCA 338
Dd	883	AGGAAGAGGTAAACACATCATCGACGCGGTGAAGCATGGCCACGCGCACCGGAGCGGTA 824
OY	339	GCGAATATCCGCTGACGGGGTGAACAGCCCTCCGCCAACGCCCAACCGGACCGCCCT 398
Dd	823	GGGAGTACCCGGTTGTTCGATGAAAACGCCGCGCGCTGTAATCCGTACAGGCTTCGGTT 764
OY	399	GCGGTGTTCCGCGCAGAAAGCCTTATGGGTCATGGGCGCAGGCGCATGGGCGAGATGCCCC 458
Dd	763	GC-TGTTGGCAAAACTGACGATTAATCGCCGTTAAAGTAATGGGCAATGCAACCTGTTTC- 706
OY	459	TTTTGGCGCGGATCTCTGCCCCGGTCCAGCCCGCCGCTGCGGSCATATGTCAGCGACGCT 518
Dd	705	-TTCCCGGACGACGTCTGTAAACGGCCCAACCTGTGCGGCGACATTAATCGGAATGTTCT 647
OY	519	GCGCCAGCGCCCATCTGTCAGATTCGCGCGCGCTGCTAGACGCGGTATCTTCGATCAGGA 578
Dd	646	GACGCGCCGTTCCGCTCGAAGATTAGCCTTGTCAATGATGATGATGATGATGATGATGATG 587
OY	579	TGCGGGTGGACTGAAGGCGACGATAGAAAGCGGTACCCGTCATCTTCGCGGAACGG 638
Dd	586	GTGGCGTTGCGAAAAGCGCAGGGTATTAACAAAGCGGTACCAATTTGTGATGACCG 527
OY	639	TGCGGTTCATATCATCGGGGCGTTCGACGCGATGGGGGCGGTTCGATCTCGACGC 698
Dd	526	TGCGATTCATATANTTCGTGTACATTAACATGCGGCGCTCAGTTGCACTCTTGAC 467
OY	699	CCAGGAATTTCTGAAACCCAACGCTCAGGTGCGGGGTCTCGACGCGCACACGAGCGGTGCA 758
Dd	466	CGATTAATGCTTGGATCTTAACGGGTATGCAATCAGGGGTGTAACCCCGTTCGTGGA 407
OY	759	TGACGACGCGACGCTCGATCCGCGAAGCCTTCGTCAGCGTGGCGCGGATTCGTACGCG 818

Db 406 TCACGTACTGCGATGATATCCGGCATCCGCTACAGTGAACCGATTCAGCATGAACGG 347
Qy 819 TCGGACATGCGATTCACCGGAGATGACACCCCTGACGACCGGATGAGCGG---- 874
Db 346 CTGAACCGCGGATGACCGATTAATGTTGTCMAAATGTTGCGGAGTATCCCGCGCA 287
Qy 875 -----CCCGCTCGATGACGACCATAGCTGTCTGTCAGGCGCGCGAATGTGCGGAAACG 929
Db 286 AATGCGGGGAGGATGACGACGATAGTACACTGTTTCAATGCGACGCGCTTGCGGGGAAAC 227
Qy 930 CGACTCTTCGATCCGTCATTCGCCCGGAGAAATGGCGGACAGCGCGCGCAATTCGG 989
Db 226 GAACCTGTGATGCGGCGCAAGTGAAGCAACGCGCGCGCTATCCAGCGATGCTGATCA 167
Qy 990 GCGAAGATCCGCTGTGCGGACGACGATGCTGTCTGTCAGGCGCGCGGCGCGCGCT 1049
Db 166 GCGTTAAATCTCTTGTGAAAGACAGGATAGTGTCTTCCCGCTCGAGACCGCGCT 107
Qy 1050 CGAGCATCAAGATGCGGATCCGCTGTGCGGTCGCGAAGCGAAGCGCATGCGGAC 1109
Db 106 CATAAGCAAGATCCGATATCCGATATCCGATGCTGTGTAAGCGGAGCGCATGAAGCGCAT 47
Qy 1110 CGGACAGCGCGCGCGCGCGCATCAAGCATCATGG 1144
Db 46 TAGCCAGACCGGACCGACCGAATGATCATAG 12

RESULT 15
ADQ77203/c
ID ADQ77203 standard; DNA; 1149 BP.
XX
AC ADQ77203;
XX
DT 23-SEP-2004 (first entry)
XX
DE Nucleotide sequence of Pantoea stewartii CrtY protein.
XX
KM carotenoid: carotenoid enzymatic biosynthetic pathway; dxs gene;
D-1-deoxyxylulose 5-phosphate gene; idi gene; ygbB gene;
KM yler gene; oligo-ribonuclease gene; dxs gene; dxt gene;
DOXP reductoisomerase; ygbp gene; ispp gene; ycbB gene; ispb gene;
KM ygbp gene; ispp gene; lytB gene; ispp gene; idi gene;
KM isopentenyl diphosphate isomerase; ispp gene;
KM farnesyl pyrophosphate synthase; ispp gene;
octaprenyl diphosphate synthase; crtB gene;
KM geranylgeranyl pyrophosphate synthase; crtB gene; phytoene synthase;
KM crtI gene; phytoene dehydrogenase; crtX gene;
KM zeaxanthin glucosyl transferase; crtZ gene; beta-carotene hydroxylase;
KM crtY gene; lycopenase cyclase; enzyme; gene; ds.
OS Pantoea stewartii.
XX
XX
XX Key Location/Qualifiers
FT 1.1149
FT /*tag= a
FT /product= "CrtY protein"
XX
XX MO2004056975-A2.
XX
XX 08-JUL-2004.
XX
XX 19-DEC-2003; 2003MO-US041812.
XX
XX 19-DEC-2002; 2002US-0434618P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
XX Cheng Q, Rouviere PE, Suh W;
XX
XX MPI; 2004-525437/50.
XX
XX P-PSDB; ADQ77204.
XX
XX Novel carotenoid overproducing bacteria such as Escherichia coli or
XX
XX

PT Flavobacterium, comprising genes encoding functional carotenoid enzymatic
PT biosynthetic pathway, useful for producing carotenoids such as
PT zeaxanthin and beta-carotene.
PS Example 4; SEQ ID NO 5; 137pp; English.
XX
XX The specification describes a carotenoid overproducing bacteria, which
CC comprise genes encoding a functional carotenoid enzymatic biosynthetic
CC pathway where the dxs (D-1-deoxyxylulose 5-phosphate gene), idi, ygbB
CC and ispb genes are overexpressed, and the yler gene (oligo-ribonuclease
CC gene) is down regulated. In bacteria of the invention, the carotenoid
CC enzymatic biosynthetic pathway consists of the genes dxs, dxt (DOXP
CC reductoisomerase), ygbp (also known as ispp), ycbB (also known as ispb),
CC ygbB (also known as ispp), lytB (also known as ispp), idi (isopentenyl
CC diphosphate isomerase), ispa (farnesyl pyrophosphate synthase), ispp
CC (octaprenyl diphosphate synthase), crtB (geranylgeranyl pyrophosphate
CC synthase), crtB (phytoene synthase), crtI (phytoene dehydrogenase), crtX
CC (zeaxanthin glucosyl transferase), crtZ (beta-carotene hydroxylase) and
CC crtY (lycopenase cyclase). The lytB and dxt gene are optionally
CC overexpressed. Carotenoid overproducing bacteria of the invention are
CC useful for producing carotenoids. The present sequence encodes a CrtY
CC protein.
SQ Sequence 1149 BP; 228 A; 325 C; 315 G; 281 T; 0 U; 0 Other;
XX
XX Query Match 18.3%; Score 209.8; DB 12; Length 1149;
Best Local Similarity 51.8%; Pred. No. 2.5e-30;
Matches 578; Conservative 0; Mismatches 522; Indels 15; Gaps 4;
Qy 42 GCACGCGCTGCGACACGCGGATGCGGCGCGCTCGGTCAGATGCGAACCCTGCGGCA 101
Db 1123 GCATGCGCGCGGAAACGGGAGAGCGGCGCTTCCGCTGAGATGCGTACCGCATCGCTCA 1064
Qy 102 ATGTACGCGCGCGCGGATGAAAGCTCGATCAAGCGGCTGCGGAGCGGTAACCGCT 161
Db 1063 CCGTGAATTTTCCCGATTAAGCGGCAATCAATCTCGGATGAGCATGAACCGCT 1004
Qy 162 GCACGCGGATGAGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 221
Db 1003 GCATCAACGCGGACGCGTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 944
Qy 222 GCAGAAAGCGGCGGCGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 278
Db 943 GGAAGAAACCGCTGTGCTCAGCAAGTGTGCGGCAAGTGAAGCAATGCTGTGTAACAG 884
Qy 279 CGAGCGGCTGCTGAGTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 338
Db 883 AGGAAGAGGTAACACATCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 824
Qy 339 GCGAATATCCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 398
Db 823 GCGAGTAGCGGCTGTGCGATTAACGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 764
Qy 399 GCGCGTGTGCGGCGGAGGAGCTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 458
Db 763 GC-TGTGCGAAGATGAGATTAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 706
Qy 459 TTTGCGCGCGATCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 518
Db 705 -TTCCCGAGCAACGCTGTAAACGCGCAACCTGTGCGGCGGCGGCGGCGGCGGCGGCG 647
Qy 519 GCGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 578
Db 646 GACGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 587
Qy 579 TCGGCGTGGAGCTGAAGGCGGAGAGATGATGAAGCGGATCCGTCATCTGCGGAACGG 638
Db 586 GTGCGGTTGCGAAGGCGGAGGATTAACAAAGGCGGATTTGCTATGACCG 527
Qy 639 TCGCGTCAATGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 698
Db 526 TCGCATCAATGATGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 467

QY	699	CCAGAAATTTCTGGAAACCCAGGGTACAGGTGCGGGGTCTCGACGGCACACACGGGCGTCGA	758
Db	466	CGATTAATGCTCGGAAATCTTACGGGTATGCAGAATACAGGCGTTAAACCCGTCGTCGA	407
QY	759	TCACGACAGGACGCTCGATTCGCGACAGCCGTCGTCAGGTCGCGCCGGTATCGTCCACG	818
Db	406	TCACGTACTGGCATGAATTAATTCGGGCATCCGCTTAATCGAACCGATTACGATGAACGG	347
QY	819	TCGCGACATGGGTATTCACCGCGAGATGCAACCTGACGACGGCCCATCGACGGCG----	874
Db	346	CTGAAACCGCGGTATGACGCATTAATGTTGTCTCAAACTGTTGCGCGAGTATCCCGGCGA	287
QY	875	-----CCGCGCTCGATCGAGCCATAGACCTGTCGTACAGCGCGGCGGAATGTCGGGAAACG	929
Db	286	AATGCGGGAAGTCAACGAGTAGTAGCCACTGTTTACATGGCGAOCGGGTTTGGGGGAAAC	227
QY	930	CGACCTCTGATCCGTCATTGCGCGCGACGAATGGCGGACAGAGGCGCGCCAGCCATTGCG	989
Db	226	GAACCTGGTAGTCGGGCGACGATGACACAAAGCGGCGTATTCAGCGATGCTGAATTCA	167
QY	990	GCGAAGATTCGTGTCGTGCGGACGACAGGTGTGCTGATTCGAGGGGCGGACCGCGCGT	1049
Db	166	GCGTTAAATCTCTTTCGTGAAGGACCAAGGATGTTTCTCTTCGCGCTCAGAACCCGCTT	107
QY	1050	CGAGCATCAGATGGCGCATTCGGTCTGCGGTGCGCAACGACGACCGATCGACGCGAC	1109
Db	106	CAATTAAGCAAGATCCGATATTCGGGATGCTGTGCTGAAGCGCGAGCGCGATTAGGCGAT	47
QY	1110	CGGACAGCCCGCGCGCGCGGATCAGACGATCATGG	1144
Db	46	TAGCCAGACGGCACCGACGAAATGAATCATATG	12

Search completed: June 21, 2005, 20:32:40
Job time : 756.928 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 18:02:15 ; Search time 5324.07 Seconds
(without alignments)
10457.239 Million cell updates/sec

Title: US-10-695-980-1_COPY_5794_6942

Perfect score: 1149
Sequence: 1 tcaagctctctccctgcagca.....atcagcagatcatgctcat 1149

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	1149	100.0	1149	6 A58567	A58567 Sequence 8
C 2	1149	100.0	1149	6 A84704	A84704 Sequence 14
C 3	1149	100.0	1149	6 AR103082	AR103082 Sequence
C 4	1149	100.0	1149	6 AR139698	AR139698 Sequence
C 5	1149	100.0	1149	6 AR391877	AR391877 Sequence
C 6	1149	100.0	5188	1 PAU62808	U62808 Flavobacter
C 7	1149	100.0	8625	6 AR159831	AR159831 Sequence
C 8	1149	100.0	8625	6 AR452005	AR452005 Sequence
C 9	1149	100.0	11233	6 AR169852	AR169852 Sequence
C 10	1149	100.0	11233	6 AR452026	AR452026 Sequence
C 11	641.8	55.9	1161	6 AX429765	AX429765 Sequence
C 12	640.8	55.8	1161	6 AR041163	AR041163 Sequence
C 13	640.8	55.8	1161	6 AR082017	AR082017 Sequence
C 14	640.8	55.8	1161	6 E22378	E22378 beta-Carote
C 15	640.8	55.8	1161	6 AR041164	AR041164 Sequence
C 16	640.8	55.8	2886	6 AR082018	AR082018 Sequence
C 17	640.8	55.8	2886	6 AX429767	AX429767 Sequence
C 18	640.8	55.8	5173	1 ATUCRTWA	D58420 Paracoccus
C 19	624.8	54.4	6335	1 PMCRWZYI	Y15112 Paracoccus

C 20	395.4	34.4	5397	12 AY605097	AY605097 Synthetic
C 21	296.8	25.8	3801	1 AF408848	AF408848 Xanthobac
C 22	236.4	20.6	6635	1 AF218415	AF218415 Bradyrhiz
C 23	216	18.8	6815	1 AB076662	AB076662 Pantoea a
C 24	213.4	18.6	12753	1 ERWCRTS	M87280 Pantoea ag
C 25	209.8	18.3	1149	6 AX394986	AX394986 Sequence
C 26	209.8	18.3	6586	1 AY166713	AY166713 Pantoea s
C 27	200	17.4	1235	6 I72656	I72656 Sequence 13
C 28	198.4	17.3	1235	6 I72657	I72657 Sequence 15
C 29	174.8	15.2	1149	6 E22382	E22382 beta-Carote
C 30	174.8	15.2	1149	6 E22382	E22382 beta-Carote
C 31	174.8	15.2	4624	6 CQ793208	CQ793208 Sequence
C 32	174.8	15.2	6918	6 ERWCRT	D90087 Erythrina ure
C 33	174.8	15.2	6918	6 I12910	I12910 Sequence 13
C 34	173.2	15.1	6965	1 ERWCRTA	M90698 Pantoea ag
C 35	138.8	12.1	1631	1 APACRTWB	D58422 Alcaligenes
C 36	138.8	12.1	1631	6 AR041167	AR041167 Sequence
C 37	138.8	12.1	1631	6 AR082021	AR082021 Sequence
C 38	138.8	12.1	1631	6 CQ788337	CQ788337 Sequence
C 39	138.8	12.1	1631	6 CQ793304	CQ793304 Sequence
C 40	138.8	12.1	1631	6 CQ793485	CQ793485 Sequence
C 41	138.8	12.1	1631	6 AX429772	AX429772 Sequence
C 42	107.2	9.3	349723	1 BX842650	BX842650 Bdelioid
C 43	99.6	8.7	349107	1 BX571873	BX571873 Photobact
C 44	93.8	8.2	1305	1 D83513	D83513 Erythrobact
C 45	93.8	8.2	1305	6 E11104	E11104 Gene coding

ALIGNMENTS

RESULT 1	A58567/c	1149 bp	DNA	linear	PAT 06-MAR-1998
LOCUS	A58567	Sequence 8 from Patent EP0747483.			
DEFINITION	A58567				
ACCESSION	A58567				
VERSION	A58567.1	GI:3714158			
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1				
AUTHORS	Hohmann,H., Pasamonnes,L., Teseier,M. and Van,L.A.				
TITLE	Recombinative carotenoid production				
JOURNAL	Patent: EP 0747483-A 8 11-DEC-1996;				
COMMENT	HOFFMANN LA ROCHE (CH)				
FEATURES	Other publication JP 9023888 970128.				
source	1. 1149				
ORIGIN	/organism="unidentified"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32644"				
Query Match	100.0%	Score 1149;	DB 6;	Length 1149;	
Best Local Similarity	100.0%	Pred. No. 3.5e-141;			
Matches 1149;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	TCATGCTCTCTCTGACGAGGGGCGTTGCGGACGACGCGCTGCGACAGCGG	60		
DB	1149	TCATGCTCTCTCTGACGAGGGGCGTTGCGGACGACGCGCTGCGACAGCGG	1090		
QY	61	AATGGCGGGCGTCCGCTGACGATGCGAAGCCGCTGCGCAATGTCAGGCGCCGCGCAT	120		
DB	1089	AATGGCGGGCGTCCGCTGACGATGCGAAGCCGCTGCGCAATGTCAGGCGCCGCGCAT	1030		
QY	121	GAAGCGCTGATGAGCGGCTGCGGACGCGGTGAGAACCGCTGACAGGCGATAGGACG	180		
DB	1029	GAAGCGCTGATGAGCGGCTGCGGACGCGGTGAGAACCGCTGACAGGCGATAGGACG	970		
QY	181	GTCGCGCGGCGACGCGCGGACGATCGGCTTCAGCAGCCGCGAGAGCGGTCGCGATC	240		
DB	969	GTCGCGCGGCGACGCGCGGACGATCGGCTTCAGCAGCCGCGAGAGCGGTCGCGATC	910		

OY	241	CGCGGATCGATGATGAGCCGACGCGCGGACCGCGGACGAGCGGCGGACGCGGCGTCTGATGAGTCCGG	300
Dp	909	CGCGCGATTCGATGATGACCCAGACCGCGGACCGCGCGGACCGGCGGACCGCGGCTGTCAAGTCCGG	850
OY	301	CGCCCGCATGCGATTCGCGGACCTTCGCGCGGACATAGGGCAGCGAATATCCGCTGACCGGAGTG	360
Dp	849	CGCCGCGAATGGCATTCGCGGACCTTCGCGCGGACATAGGGCAGCGAATATCCGCTGACCGGAGTG	790
OY	361	GAACAGCCCTGACCCCGACGCGGACCGCGCGGACCGCGGCTTCGCGGCTGTGCGCGCAGAAAGCC	420
Dp	789	GAACAGCCCTGACCCCGACGCGGACCGCGCGGACCGCGGCTTCGCGGCTGTGCGCGCAGAAAGCC	730
OY	421	TATGAGCTCATATGAGGCGAGCGGACATGAGGACGATGACCCCTTTCGCGCGCGCATCTCTGACCC	480
Dp	729	TATGAGCTCATATGAGGCGAGCGGACATGAGGACGATGACCCCTTTCGCGCGCGCATCTCTGACCC	670
OY	481	GATCCAGCCCGCGCTTCGCGCGGACATAGTTCAGCGGACCGCTTCGCGCGGACCGCGCATTCGTCCAG	540
Dp	669	GATCCAGCCCGCGCTTCGCGCGGACATAGTTCAGCGGACCGCTTCGCGCGGACCGCGCATTCGTCCAG	610
OY	541	ATCCGCGCGCTTCGCTGTAGCGCGTATCTTCGATTCAGGATTCGCGGCTGCGGACTGAAAGGCGAG	600
Dp	609	ATCCGCGCGCTTCGCTGTAGCGCGTATCTTCGATTCAGGATTCGCGGCTGCGGACTGAAAGGCGAG	550
OY	601	CAGATAGATGGAAGCGGCTACCGGTCATCTGCGGGAACCGGCTCGGTCATAGATATGCGGGCG	660
Dp	549	CAGATAGATGGAAGCGGCTACCGGTCATCTGCGGGAACCGGCTCGGTCATAGATATGCGGGCG	490
OY	661	CTCGACGCGCATAGGCGGGGCGTCGGATCTCGATCTCGACGCGCCACGAATTTCTGGAACCCGAC	720
Dp	489	CTCGACGCGCATAGGCGGGGCGTCGGATCTCGATCTCGACGCGCCACGAATTTCTGGAACCCGAC	430
OY	721	GGTCAAGGTTCGGGAGTTCGACCGGCAACACGGGCGTTCGATCACGCAAGGCAAGCTTCGATCCG	780
Dp	429	GGTCAAGGTTCGGGAGTTCGACCGGCAACACGGGCGTTCGATCACGCAAGGCAAGCTTCGATCCG	370
OY	781	CGAGCGCGTCCGTCACGCGTTCGCGCGGCTATTCGTCACAGCGTTCGCGGAATCGTATTCACACG	840
Dp	369	CGAGCGCGTCCGTCACGCGTTCGCGCGGCTATTCGTCACAGCGTTCGCGGAATCGTATTCACACG	310
OY	841	CAGATTCGACACCTTCGACGAGCCCGCATAGCGCGCGCGCTTCGATTCGAGCCCATAGCTGT	900
Dp	309	CAGATTCGACACCTTCGACGAGCCCGCATAGCGCGCGCGCGCTTCGATTCGAGCCCATAGCTGT	250
OY	901	CGTCAGGCGGCGCGAATGCTCGGGAACCGCGACCTTCGTATTCGTCATTTCGCGCGGACG	960
Dp	249	CGTCAGGCGGCGCGAATGCTCGGGAACCGCGACCTTCGTATTCGTCATTTCGCGCGGACG	190
OY	961	AATGGGCGACAGCGCGCGCGCATTCGCGGCGGAAGATCCGCTGTGTGGCAGAGACCAAGT	1020
Dp	189	AATGGGCGACAGCGCGCGCGCATTCGCGGCGGAAGATCCGCTGTGTGGCAGAGACCAAGT	130
OY	1021	GTCGCTGTTCGAGGAGGCGGACCGCGCGCTTCGAGCATTCAGATGCGCGCATCCGCTCTCGG	1080
Dp	129	GTCGCTGTTCGAGGAGGCGGACCGCGCGCTTCGAGCATTCAGATGCGCGCATCCGCTCTCGG	70
OY	1081	GTCGCGAACGCGAAGCGCGGATCAGCGCACCGGACAGCCCGCGCGCGCATCAGCAGATC	1140
Dp	69	GTCGCGAACGCGAAGCGCGGATCAGCGCACCGGACAGCCCGCGCGCGCATCAGCAGATC	10
OY	1141	ATGGCTCAT 1149	
Dp	9	ATGGCTCAT 1	

RESULT 2
A84704/C
LOCUS A84704 1149 bp DNA linear
DEFINITION Sequence 14 from Patent EP0872554.
ACCESSION A84704
VERSION A84704.1 GI:6733573
KEYWORDS .

SOURCE	unidentified
ORGANISM	unidentified
REFERENCE	unclassified.
AUTHORS	1 (bases 1 to 1149)
TITLE	Pasamonis, L. and Tsygankov, Y. S.
JOURNAL	Improved fermentative carotenoid production
	Patent: EP 0872554-A 14 21-OCT-1998;
	HOFFMANN LA ROCHE (CH)
FEATURES	Location/Qualifiers
SOURCE	1..1149
	/organism="unidentified"
	/mol_type="unassigned DNA"
	/db_xref="Caxon:32644"
ORIGIN	

Query Match	100.0%	Score 1149;	DB 6;	Length 1149;
Best Local Similarity	100.0%	Pred. No.3.5e-141;		
Matches 1149;	0;	Mismatches 0;	Indels 0;	Gaps 0

OY	1	CAATGCTCTCCATGCAAGCAGGGGGCGTTGGGGGAGGAGGCAACCGGCTGCAACAGCGG	60
Db	1149	TCATGCTCTCTCTGACAGAGGGGGCGTTGGGGGAGGAGGCAACCGGCTGCAACAGCGG	1099
OY	61	AATGGGCGGGGCGTCCGATGACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCCGGCATA	120
Db	1089	AATGGGCGGGGCGTCCGATGACGATGCGAAGCCGGTCCGCAATGTCAGGCGCCCGGCATA	103
OY	121	GAAAGCGCTCGATCAGCGGCTGCGGAGGCGGTAGAACCTGCTGCAAGCAGGCGATAGCCAGC	180
Db	1029	GAAACGCGCTCGATCAGCGGCTGCGGAGGCGGTAGAACCTGCTGCAAGCAGGCGATAGCCAGC	970
OY	181	GTCCGGGCGGGGAGGCGGGGAAACACATCCGGTTACGACGCGCGAGGAAAGCGGTGGCGATC	240
Db	969	GTCCGGGCGGGGAGGCGGGGAAACAGATCCGGTTACGACGCGCGAGGAAAGCGGTGGCGATC	910
OY	241	CGCGCGATTCGATGCGGCCCAAGCGCGCACCGCGCGACGAGCGGCGGTCTCAGGTCGCG	300
Db	909	CGCGCGATTCGATGCGGCCCAAGCGCGCACCGCGCGACGAGCGGCGGTCTCAGGTCGCG	850
OY	301	CGCGCGATGCGATCCGCGACCTTGCGGGCACTAGGGGACGCAATATCCGGTGA CGGGGTG	360
Db	849	CGCGCGATGCGATCCGCGACCTTGCGGGCACTAGGGGACGCAATATCCGGTGA CGGGGTG	790
OY	361	GAAACGCGCTTGCCTCCCAAGCCCAACCGGCAACCGCCCTTGCGGCGGTGTGGCGCAGAAACC	420
Db	789	GAAACGCGCTTGCCTCCCAAGCCCAACCGGCAACCGCCCTTGCGGCGGTGTGGCGCAGAAACC	730
OY	421	TATGACGTCATGAGGACAGCGGATGGGAGATGGCCCTTTCGCGCGGCATCTCTTGCC	480
Db	729	TATGACGTCATGAGGACAGCGGATGGGAGATGGCCCTTTCGCGCGGCATCTCTTGCC	670
OY	481	GGTCCAGCCCGCGCTGGGCGGATGTCAGGCAAGCCTTGCAGCCAGCGCGGCATGTCAG	540
Db	669	GGTCCAGCCCGCGCTGGGCGGATGTCAGGCAAGCCTTGCAGCCAGCGCGGCATGTCAG	610
OY	541	ATCGCCCGCGCTGCTGTAGCCGATATCTCGATCAGGATGCGGGTGGGACTTGAAAGGAG	600
Db	609	ATCCCGCGCGCTGCTGTAGCCGATATCTCGATCAGGATGCGGGTGGGACTTGAAAGGAG	550
OY	601	CAGATAGTGAAGGAGTACCCTGCATCTCGGAGAAAGGTGCGCTCATGATCATCGGGCG	660
Db	549	CAGATAGTGAAGGAGTACCCTGCATCTCGGAGAAAGGTGCGCTCATGATCATCGGGCG	490
OY	661	CTCAGACGCATGAGGGGGGCGTGCCTCGATCTCGACGCCCAAGATTTCTGGAAACCCAG	720
Db	489	CTCAGACGCATGAGGGGGGCGTGCCTCGATCTCGACGCCCAAGATTTCTGGAAACCCAG	430
OY	721	GGTCAAGGTGCGGGGCTTCGACGAGCAACCGGGCGTCGATCAACGACGAGCCTTCGATCCG	780
Db	429	GGTCAAGGTGCGGGGCTTCGACGAGCAACCGGGCGTCGATCAACGACGAGCCTTCGATCCG	370
OY	781	CGAGCCGTCCGTCAAGCGTCGCGCGGATATGTCGAGGTCGCCGACATGCGATTTCCACCG	840

Db	849	CGCCGCGATGGCATCCCGGACCTGCGGGGCGATAGGGGACGAGGAATATCCGGTAGACGGGGTG	790
Qy	361	GAACAGGCTTGGCCCCCAGCCCAACCGGACCGGACCGCCCTTGGCGGTGTGGCGCCAGAAAGCC	420
Db	789	GAACAGGCTTGGCCCCCAGCCCAACCGGACCGGACCGCCCTTGGCGGTGTGGCGCCAGAAAGCC	730
Qy	421	TATGGCGTCAATGGGCGCAGCGCGAATGGGCGAGAAATGCCCTTTCGGCGCGGCATCTCTGCGC	480
Db	729	TATGGCGTCAATGGGCGCAGCGCGAATGGGCGAGAAATGCCCTTTCGGCGCGGCATCTCTGCGC	670
Qy	481	GGTCACAGCCCGCGCTTGGCGGCGATAGTTCAGAGGACGCGCTTGGCGCGACGCGCGCATCTGTCAG	540
Db	669	GGTCACAGCCCGCGCTTGGCGGCGATAGTTCAGAGGACGCGCTTGGCGCGACGCGCGCATCTGTCAG	610
Qy	541	ATCGCCCGCGTCTGCTGTAGCGCGGTATCTCGATCAGGATGGCGGGTGGGACTGAAGGCGAG	600
Db	609	ATCGCCCGCGTCTGCTGTAGCGCGGTATCTCGATCAGGATGGCGGGTGGGACTGAAGGCGAG	550
Qy	601	CAGATTAAGTAAAGCGGTAAACCGGTTCATCTGTGGGAAAGCGTGGCGTTCATTAATTCGGGCG	660
Db	549	CAGATTAAGTAAAGCGGTAAACCGGTTCATCTGTGGGAAAGCGTGGCGTTCATTAATTCGGGCG	490
Qy	661	CTGACGCGCATTGGGGGGGCGTCCGCTCGATCTCGACGCGCCAGAAATTTCTTGGAAACCCAG	720
Db	489	CTGACGCGCATTGGGGGGGCGTCCGCTCGATCTCGACGCGCCAGAAATTTCTTGGAAACCCAG	430
Qy	721	GGTCAGGTTGGCGGGGTCTCGACGCGACCAAGGCGCTCGATCAACGAGGACGCTTCGATCCG	780
Db	429	GGTCAGGTTGGCGGGGTCTCGACGCGACCAAGGCGCGTTCGATCAACGAGGACGCTTCGATCCG	370
Qy	781	CGAGCCGTCGTCAGCGCTCGGCGCGGTATGTCTCAAGCGTGGCGCATATGCTATTCACCG	840
Db	369	CGAGCCGTCGTCAGCGCTCGGCGCGGTATGTCTCAAGCGTGGCGCATATGCTATTCACCG	310
Qy	841	CAGATTCAGAACCCCTGACAGGACGCCCGATCAGCGGCGCGCGCGCTCGATTCAGAGCATATGAGCTGT	900
Db	309	CAGATTCAGAACCCCTGACAGGACGCCCGATCAGCGGCGCGCGCGCTCGATTCAGAGCATATGAGCTGT	250
Qy	901	CGTCAGGCGGCGCGCATATGTCGGGAAACGCGCACTCTCGATTCGTCATTCGCGCGGACAG	960
Db	249	CGTCAGGCGGCGCGCATATGTCGGGAAACGCGCACTCTCGATTCGTCATTCGCGCGGACAG	190
Qy	961	AATGGGCGACAGCGCGCGCCAGCCATTTCGGGCGAAAGATTCGGTGTCTGTGGCAGGACAGGT	1020
Db	189	AATGGGCGACAGCGCGCGCCAGCCATTTCGGGCGAAAGATTCGGTGTCTGTGGCAGGACAGGT	130
Qy	1021	GTCGTGATTCGAGGGGGCGCGACCGCGCGTTCGAGATACATACATGGCGCCATTCGGTCTGGCG	1080
Db	129	GTCGTGATTCGAGGGGGCGCGACCGCGCGTTCGAGATACATACATGGCGCCATTCGGTCTGGCG	70
Qy	1081	GTCGCGAAGCGCAAGCGCGATTCAGCGACCGGACGCGCGCGCGCGATTCAGCAGATC	1140
Db	69	GTCGCGAAGCGCAAGCGCGATTCAGCGACCGGACGCGCGCGCGCGATTCAGCAGATC	10
Qy	1141	ATGGCTCAT	1149
Db	9	ATGGCTCAT	1
RESULT 4			
ARI39698/c			
LOCUS			
ARI39698			
DEFINITION			
Sequence 8 from patent US 6207409.			
ACCESSION			
ARI39698			
VERSION			
ARI39698.1			
GI:14482194			
KEYWORDS			
Unknown.			
SOURCE			
Unknown.			
ORGANISM			
Unclassified.			
1 (bases 1 to 1149)			
REFERENCE			
AUTHORS			
Hohmann,H.-P., Pasamontes,L., Tessier,M. and van Loon,A.			
TITLE			
Fermentative carotenoid production			
Patent: US 6207409-A 8 27-MAR-2001;			
JOURNAL			

FEATURES

Location/Qualifiers
1. 1149
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1149; DB 6; Length 1149;
Best Local Similarity 100.0%; Pred. No. 3.5e-141;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCATGCTCTCTCTGACGAGGGGGGCTTCGGGACGAGCGGACGAGCGGCTGAGACGCGG 60
1149 TCATGCTCTCTCTCTGACGAGGGGGGCTTCGGGACGAGCGGACGAGCGGCTGAGACGCGG 1090
61 AATGGCGGGCGTCCGGTACGATGCGAAGCGGTCGGCCAAATGTCAGCGCGCCGCATATA 120
1089 AATGGCGGGCGTCCGGTACGATGCGAAGCGGTCGGCCAAATGTCAGCGCGCCGCATATA 1030
121 GAAGCGCTCTGATCAGCGGCTGCGGACGCGGTTAGAACCCCTGACAGCAGCGATAGCAGC 180
1029 GAAGCGCTCTGATCAGCGGCTGCGGACGCGGTTAGAACCCCTGACAGCAGCGATAGCAGC 970
181 GTCCGGCGGGGACGCGCGGACAGCATCCGGTTTCAGCAGCCGACAGAACCGGTCGGGATC 240
969 GTCCGGCGGGGACGCGCGGACAGCATCCGGTTTCAGCAGCCGACAGAACCGGTCGGGATC 910
241 CGCGCGATGATGAGCCGACCGCGACCGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 300
909 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 850
301 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 360
849 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 790
361 GAACGCGCTTGGCCCCCGACCGGACCGCGGCGGACCGGCGGTCGTCAGGTCGCGCAGAACGC 420
789 GAACGCGCTTGGCCCCCGACCGGACCGCGGCGGACCGGCGGTCGTCAGGTCGCGCAGAACGC 730
421 TATGGCGTATGAGGCGGACGCGGATGAGGCGGATGCGCCCTTTGCGCGCGGATCTCTGCCC 480
729 TATGGCGTATGAGGCGGACGCGGATGAGGCGGATGCGCCCTTTGCGCGCGGATCTCTGCCC 670
481 GGTCCAGCCCCGCTGGCGGACATAGTCCAGCGACCGCTTCGCGGACGCGGTCGTCAGGTCGCG 540
669 GGTCCAGCCCCGCTGGCGGACATAGTCCAGCGACCGCTTCGCGGACGCGGTCGTCAGGTCGCG 610
541 ATCCGCGCGGCTGCTGTAGCGGATCTCTGATCAGAGGCGGCGGTCGCGGTCGCGGTCGCGG 600
609 ATCCGCGCGGCTGCTGTAGCGGATCTCTGATCAGAGGCGGCGGTCGCGGTCGCGGTCGCGG 550
601 CAGATAGATGAGCGGCTACCCGTCATCTGCGGAAAGGTCGCGTCATGATCATCGGCGG 660
549 CAGATAGATGAGCGGCTACCCGTCATCTGCGGAAAGGTCGCGTCATGATCATCGGCGG 490
661 CTCGACGCGCATGAGGCGGCGTCCGTCGATCTTCGACGCGCCACGAAATTTCTGAAACCCAC 720
489 CTCGACGCGCATGAGGCGGCGTCCGTCGATCTTCGACGCGCCACGAAATTTCTGAAACCCAC 430
721 GGTTCAGGTCGCGGCTTCGACGCGGACGAGGCGGTCGATCAGAGGCGGCTTCATTCG 780
429 GGTTCAGGTCGCGGCTTCGACGCGGACGAGGCGGTCGATCAGAGGCGGCTTCATTCG 370
781 CGAGCGCTCCGTCAGCGGCTGCGGATCTGTCAGCGGTCGCGGACGATGCGGATTCGACCG 840
369 CGAGCGCTCCGTCAGCGGCTGCGGATCTGTCAGCGGTCGCGGACGATGCGGATTCGACCG 310
841 CAGATGACACCTCTGACGAGCGCGGATCAGCGGCGCGCGCTTCGATGAGCCATAGCTGT 900
309 CAGATGACACCTCTGACGAGCGCGGATCAGCGGCGCGCGCTTCGATGAGCCATAGCTGT 250
901 CGTCAGGCGGCGGATGAGTCGCGGAAACGCGACCTCTGATCCGTCATTCGCGCGGAGG 960
249 CGTCAGGCGGCGGATGAGTCGCGGAAACGCGACCTCTGATCCGTCATTCGCGCGGAGG 190

QY 961 AATGGCGCAGAGCGCGCCAGCCATTCGGGCGAAAGATCCGTGTCGGCAGAGCAGGT 1020
DB 189 AATGGCGCAGAGCGCGCCAGCCATTCGGGCGAAAGATCCGTGTCGGCAGAGCAGGT 130
QY 1021 GTGCTGTGTCGAGGGGCGCGACCGGCGCTTCAGCATCAGATGCGCGGATCCGCTTCG 1080
DB 129 GTGCTGTGTCGAGGGGCGCGACCGGCGCTTCAGCATCAGATGCGCGGATCCGCTTCG 70
QY 1081 GTGCGAACGCGCAACGCGATCAGGCGACCGGACAGCCCGCGCGCGATCAGAGATC 1140
DB 69 GTGCGAACGCGCAACGCGATCAGGCGACCGGACAGCCCGCGCGCGATCAGAGATC 10
QY 1141 ATGCTCAT 1149
DB 9 ATGCTCAT 1

RESULT 5
AR391877/c 1149 bp DNA linear PAT 18-DEC-2003
LOCUS AR391877
DEFINITION Sequence 8 from patent US 6613543.
ACCESSION AR391877
VERSION AR391877.1 GI:40115617
KEYWORDS
SOURCE
ORGANISM Unknown.
UNCLASSIFIED.

REFERENCE 1 (bases 1 to 1149)
AUTHORS Hohmann, H.-P., Pasamontes, L., Tessier, M. and van Loon, A.
TITLE Fermentative carotenoid production
JOURNAL Patent: US 6613543-A 8 02-SEP-2003;
FEATURES Location/Qualifiers
1. 1149
source /organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1149; DB 6; Length 1149;
Best Local Similarity 100.0%; Pred. No. 3.5e-141;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCATGCTCTCTCTGACGAGGGGGGCTTCGGGACGAGCGGACGAGCGGCTGAGACGCGG 60
1149 TCATGCTCTCTCTGACGAGGGGGGCTTCGGGACGAGCGGACGAGCGGCTGAGACGCGG 1090
61 AATGGCGGGCGTCCGGTACGATGCGAAGCGGTCGGCCAAATGTCAGCGCGCCGCATATA 120
1089 AATGGCGGGCGTCCGGTACGATGCGAAGCGGTCGGCCAAATGTCAGCGCGCCGCATATA 1030
121 GAAGCGCTCTGATCAGCGGCTGCGGACGCGGTTAGAACCCCTGACAGCAGCGATAGCAGC 180
1029 GAAGCGCTCTGATCAGCGGCTGCGGACGCGGTTAGAACCCCTGACAGCAGCGATAGCAGC 970
181 GTCCGGCGGGGACGCGCGGACAGCATCCGGTTTCAGCAGCCGACAGAACCGGTCGGGATC 240
969 GTCCGGCGGGGACGCGCGGACAGCATCCGGTTTCAGCAGCCGACAGAACCGGTCGGGATC 910
241 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 300
909 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 850
849 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 790
301 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 360
849 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 790
DB 849 CGCGCGATGATGAGCCGACCGCGGACCGCGGCGGACGCGGCGGACGCGGTCGTCAGGTCGCG 790
QY 361 GAACGAGCCCTGCGCCGACCGCAACCGGACCGGCTTCGCGGTCGTCAGGTCGCGCAGAACGC 420
DB 789 GAACGAGCCCTGCGCCGACCGCAACCGGACCGGCTTCGCGGTCGTCAGGTCGCGCAGAACGC 730
QY 421 TATGGCGTATGAGGCGGACGCGGATGAGGCGGATGCGCCCTTTGCGCGCGGATCTCTGCCC 480
DB 729 TATGGCGTATGAGGCGGACGCGGATGAGGCGGATGCGCCCTTTGCGCGCGGATCTCTGCCC 670

QY 481 GGTCAACCCCGCTGCGGCATATGTCAGCGACGCGCTGCGCGCCCATTCGTCCAG 540
 DB 669 GGTCAACCCCGCTGCGGCATATGTCAGCGACGCGCTGCGCGCCCATTCGTCCAG 610
 QY 541 ATGCGCGCGCTGCTGTAGCGGTATCTCGATCAGATGCGGGTGGACTGAAAGGCGAG 600
 DB 609 ATGCGCGCGCTGCTGTAGCGGTATCTCGATCAGATGCGGGTGGACTGAAAGGCGAG 550
 QY 601 CAGATGATGAAAGCGGTATCCCTGATCTGCGGAAGGTCGCTCATATGATTCGGGG 660
 DB 549 CAGATGATGAAAGCGGTATCCCTGATCTGCGGAAGGTCGCTCATATGATTCGGGG 490
 QY 661 CTGCAAGCCATGGGAGGCGCTGCTGATCTGCAAGCGCCAGCAATTTCTGAAACCCAC 720
 DB 489 CTGCAAGCCATGGGAGGCGCTGCTGATCTGCAAGCGCCAGCAATTTCTGAAACCCAC 430
 QY 721 GGTCAAGTGGGAGGCTGCAAGCGCACCGGCGGTGATCAAGCAGGACGCTGATCCG 780
 DB 429 GGTCAAGTGGGAGGCTGCAAGCGCACCGGCGGTGATCAAGCAGGACGCTGATCCG 370
 QY 781 CAGAGCGTCCGTGATGAGGTGCGCGCGGTATCTGACGCTGCGACATGCGTATTCACCG 840
 DB 369 CAGAGCGTCCGTGATGAGGTGCGCGCGGTATCTGACGCTGCGACATGCGTATTCACCG 310
 QY 841 CAGATGACACCCCTGCAAGCGACCGCGCGCTGATGAGCGCATAGCCTGT 900
 DB 309 CAGATGACACCCCTGCAAGCGACCGCGCGCTGATGAGCGCATAGCCTGT 250
 QY 901 CGTCAAGCGCGCGCAATGCTCGGGAACCGGACCTTCGATCGCTGCTGCGCGGACG 960
 DB 249 CGTCAAGCGCGCGCAATGCTCGGGAACCGGACCTTCGATCGCTGCTGCGCGGACG 190
 QY 961 AATGGCGGACAGCGCGCGCAATTCGCGGCAAAATTCGCTGCTGCGGACGACGCT 1020
 DB 189 AATGGCGGACAGCGCGCGCAATTCGCGGCAAAATTCGCTGCTGCGGACGACGCT 130
 QY 1021 GTGCTGTGCGAGGCGCGGACCGCGGTGAGCATCAGATGCGCGGATCCGCTGCGG 1080
 DB 129 GTGCTGTGCGAGGCGCGGACCGCGGTGAGCATCAGATGCGCGGATCCGCTGCGG 70
 QY 1081 GTGCGCAACGCGCAAGCGCATGAGCGCACCGGACCGCGCGCGCATGAGCAGATC 1140
 DB 69 GTGCGCAACGCGCAAGCGCATGAGCGCACCGGACCGCGCGCGCATGAGCAGATC 10
 QY 1141 ATGCGTCTAT 1149
 DB 9 ATGCGTCTAT 1
 RESULT 6
 PAU62808
 LOCUS PAU62808 5188 bp DNA linear BCT 18-FEB-1997
 DEFINITION Flavobacterium ATCC21588 geranylgeranyl synthase (crtB), phycoene synthase (crtI), phycoene desaturase (crtI), lycopene cyclase (crtY), b-carotene hydroxylase (crtZ) genes, complete cds.
 ACCESSION U62808
 VERSION U62808.1 GI:1842241
 KEYWORDS Flavobacterium sp. ATCC 21588
 SOURCE Flavobacterium sp. ATCC 21588
 ORGANISM Bacteria; Bacteroidetes; Flavobacteriia; Flavobacteriales; Flavobacteriaceae; Flavobacterium.
 REFERENCE 1 (bases 1 to 5188)
 AUTHORS Pasamontes, L., Hug, D., Tessier, M., Hohmann, H. P., Schierle, J. and van Loom, A. P.
 TITLE Isolation and characterization of the carotenoid biosynthesis genes of Flavobacterium sp. strain R1534
 JOURNAL Gene 185 (1), 35-41 (1997)
 MEDLINE 97186694
 PUBMED 9034310
 REFERENCE 2 (bases 1 to 5188)
 AUTHORS Pasamontes, L.

TITLE Direct Submission
 JOURNAL Submitted (03-JUN-1996) VFGB, F. Hoffmann-La Roche Ltd., Basel, 4070, Switzerland
 FEATURES
 source location/Qualifiers
 1..5188
 /organism="Flavobacterium sp. ATCC 21588"
 /mol_type="genomic DNA"
 /strain="R1534"
 /db_xref="taxon:50286"
 95..982
 /gene="crtE"
 95..982
 /gene="crtB"
 /codon_start=1
 /transl_table=11
 /product="geranylgeranyl synthase"
 /protein_id="AAC44848.1"
 /db_xref="GI:1842242"
 /translation="MTPKQFPLRDLVEIRLAQISGQFVVSAPLGAAMSDNALSPGK RRAVLAIVMASSSGVCDAMVDACAVEMVLAALI PDMPGCMDDARTRGOPATHV AHGRRVLAIVAGTALITTEAMRIIGEARGLPPDORALVMSMRMGVPGICAGODLH APKDAAGTERBODLTGVLPAAGLEMLSTIKGIDAEFTQMAAPROGLRVRQSDYDL LDVIDKASTGKDTARDTPAEPKGLMAVGQMGDAQHTRASRAQDLBMLKRLFRG GQIDLARVLEPHDIRSA"
 complement (1890)
 /gene="crtB"
 complement (979..1890)
 /gene="crtB"
 /codon_start=1
 /transl_table=11
 /product="phytoene synthase"
 /protein_id="AAC44849.1"
 /db_xref="GI:1842243"
 /translation="WTDLTATSEALIAAGSOSFQAALIMPREGIREDTVMUYAKRHA DDVIDGQWGSAPBEGSDPQARLGARADDTLAALHEDGKSPFFALAKQVARRHPD LMPDLIEGFAMDVDAREYRSLDVLLEYSHVAGVGMMAVGVODDAVLDRACDL GLAFLQTLNARDVIDDAIIGRCYLPADMLAEGATVEGVSDALYSVIRLLDAEP YYASARQGLPHLPPRCAMSIAMALRYAIGTRIRIQGPEAVRQRI STSKAKIGILA RGLDAAASRLRGESIRSDGLWTRPRA"
 complement (1887..3371)
 /gene="crtI"
 complement (1887..3371)
 /gene="crtI"
 /codon_start=1
 /transl_table=11
 /product="phytoene desaturase"
 /protein_id="AAC44850.1"
 /db_xref="GI:1842244"
 /translation="MSSAIVIGAGFGGLAALIRLQSAIGATTIYEARDKPGRAYVMN DOGHVFDAGPTVVDPSLRBELMALSGQPMERDVTLLPVSPRYRLTMADGRSEFYND DDELIRQVASFNPADVDGRRPHDVAEEVREGVYKGLTTPRLKLGKMLNADAPALMRL QATRSVSHMVARFIODPHLRQAFSPHTLLVGNPSTSIYALITALRSGVWAKG GTNQVLAVGMVALFERLGGLTLNARVTI DVEGDATSVTLIDGROLRADVYASGDV MTSYDLGLHTRRGRTKAAIINRQMSLSFVLRHGLSKRPENLHSHSVTFQPRYKGL VNEIRFEGRLPEFMSYLPSCVTPSLAPGMSHYVALVPHLGRADVDEAAPG VAEIRFEELEBRAPIDLRKHLTVSRIFSPADPSTLSAHGSAFSEVPIILQSAWFRP HNRDAIRNFYVAGCTHPCGAGIPEVVSASATTAQVMLDLAVA"
 complement (3368..4516)
 /gene="crtY"
 complement (3368..4516)
 /gene="crtY"
 /codon_start=1
 /transl_table=11
 /product="lycopene cyclase"
 /protein_id="AAC44851.1"
 /db_xref="GI:1842245"
 /translation="MSHDLIAGAGLIALAVDRRPPARIVMLDARSGPSDOHT MSHCDTDSPEWIALSPIRGEWTDQEAEPDHRRLTTGSLIEAGALILGQGV DLRWTHVATLDDTGA TLTDGSRIRAAVCYIDARGAERTPHLTGPKFVGVEITDAPH GYERPMINDATVPQMDGYRFYTLPEFSTRLLIERTYSDDGDLDDGALAQASLDYAA RRGWGTQEMREREGILPTALAHDAIGFRDHAQGANVPGLAGLPHVPTGSLPYAQ VADAIAARDLTATASARAVGMWAIIDRADRPSLRLLNMLFRGCPDRRYRLQRFYR LPQPIERFYAGRLTLADRLRI VITGRPPI PLQAVRCMLPERPLLOERA"

gene complement(4513..5022)
/gene="crit2"
complement(4513..5022)
CDS /gene="crit2"
/codon_start=1
/transl_table=11
/product="b-carotene hydroxylase"
/protein_id="AAC44852.1"
/db_xref="GI:1842246"
translation="MSTWAILITVLITVAAEMLTAVSYVRWIMHNPGLGWMKSHHDE
DHDHLEKNDLYGVYFAVISYLPALGMSDPLAWMLAVGTVCGLIYFPLHDGLVHG
RMPFRVYPRKGYLRKYQAHNRHMHVHGENCVSPGFIWAPSDVSKLKLKSGALLK
DREGADRNT"
ORIGIN
Query Match 100.0%; Score 1149; DB 1; Length 5188;
Best Local Similarity 100.0%; Pred. No. 2.2e-141;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGCTCTCTCTGAGAGAGGAGGCGTTGCGGACAGCGCACCGGCTTGAGACAGCGG 60
DB 3368 TCATGCTCTCTCTGAGAGAGGAGGCGTTGCGGACAGCGCACCGGCTTGAGACAGCGG 3427
QY 61 AATGGGCGGCGCTCCGCTGACATGCGAAGCGGCTGCGCAATGTGAGCGCCCGGCATA 120
DB 3428 AATGGGCGGCGCTCCGCTGACATGCGAAGCGGCTGCGCAATGTGAGCGCCCGGCATA 3487
QY 121 GAAGGCGTGCATGAGGCGGCTGGGCGGAGGAGTGAACCGCTGACAGCGGATGAGGAGC 180
DB 3488 GAAGGCGTGCATGAGGCGGCTGGGCGGAGGAGTGAACCGCTGACAGCGGATGAGGAGC 3547
QY 181 GTGCGGCGGCGAGCGCGGAAACGATCCGCTTACAGCAGCGCGAGGAAACCGCTGCGATC 240
DB 3548 GTGCGGCGGCGAGCGCGGAAACGATCCGCTTACAGCAGCGCGAGGAAACCGCTGCGATC 3607
QY 241 CGCGCGATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGCTGTCAAGTCCGCG 300
DB 3608 CGCGCGATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGCTGTCAAGTCCGCG 3667
QY 301 CGCGCGATGAGCATCCGCGACCTGCGGCGATGAGGCGAGCATATCCGGTGAACGCGGAG 360
DB 3668 CGCGCGATGAGCATCCGCGACCTGCGGCGATGAGGCGAGCATATCCGGTGAACGCGGAG 3727
QY 361 GAACAGCGCTGCGCGCGGAGCGGAGCGCGGAGCGCGGCTGCGCGCGAGAGCGC 420
DB 3728 GAACAGCGCTGCGCGCGGAGCGGAGCGCGGAGCGCGGCTGCGCGCGAGAGCGC 3787
QY 421 TATGCGTATGAGGCGGAGCGGAGTGGGAGATGCCCTTTTGCGCGCATCTCTGCGCC 480
DB 3788 TATGCGTATGAGGCGGAGCGGAGTGGGAGATGCCCTTTTGCGCGCATCTCTGCGCC 3847
QY 481 GGTCCAGCGCGCGCTGGGCGGATATGTCAGCGAGCGCGTGGCGAGCGCGCATTCGCGAG 540
DB 3848 GGTCCAGCGCGCGCTGGGCGGATATGTCAGCGAGCGCGTGGCGAGCGCGCATTCGCGAG 3907
QY 541 ATGCGCGCGCTGCTGTAGCGGCTATCTCGATCAGAGATGCGGTTGGAGCTGAAGGCGAG 600
DB 3908 ATGCGCGCGCTGCTGTAGCGGCTATCTCGATCAGAGATGCGGTTGGAGCTGAAGGCGAG 3967
QY 601 CAGATGATGAAGCGGTACCGCTCATCTGCGGAAACGTTGCGGTCATATCATTCGCGGCG 660
DB 3968 CAGATGATGAAGCGGTACCGCTCATCTGCGGAAACGTTGCGGTCATATCATTCGCGGCG 4027
QY 661 CTCGAGCGCATGGGCGGCGCTGCTGCAATCTGAGCGCCACGAAATTTCTGGAAGACCGAC 720
DB 4028 CTCGAGCGCATGGGCGGCGCTGCTGCAATCTGAGCGCCACGAAATTTCTGGAAGACCGAC 4087
QY 721 GGTCAAGGTGCGGAGTTCGACAGGCAACAGGAGCGTGCATCAGCAGGAGCGCTTCGATCCG 780
DB 4088 GGTCAAGGTGCGGAGTTCGACAGGCAACAGGAGCGTGCATCAGCAGGAGCGCTTCGATCCG 4147
QY 781 CGAGCGTTCGTCAGCGCTGCGCGCGGATTCGTCACGCGTCGACATGCGTAATTCACCG 840

DB 4148 CGAGCGCTGCTGACGCGTGCAGCGCGGCTATGCTCCAGCGCTGCGACATCGTATTCACCG 4207
QY 841 CAGATCGACACCGCTGACAGCGCGCGATGAGCGCGCGCGCTCGATCGAGCCATAGCCTGT 900
DB 4208 CAGATCGACACCGCTGACAGCGCGCGATGAGCGCGCGCGCGCTCGATCGAGCCATAGCCTGT 4267
QY 901 CGTCAAGCGCGCGGATGCTGCGGAAACGCGACCTCTGATTCGCTCATTTGCGCGCGAGC 960
DB 4268 CGTCAAGCGCGCGGATGCTGCGGAAACGCGACCTCTGATTCGCTCATTTGCGCGCGAGC 4327
QY 961 AATGGGCGACAGCGCGCGAGCATTCGCGGGAAGATCCGCTGCTGCGAGGACAGAGT 1020
DB 4328 AATGGGCGACAGCGCGCGAGCATTCGCGGGAAGATCCGCTGCTGCGAGGACAGAGT 4387
QY 1021 GTGCTGCTCCAGAGGCGCGAGCGCGCGCTGACGATCAGATGCGCGCATCCGCTTCCG 1080
DB 4388 GTGCTGCTCCAGAGGCGCGAGCGCGCGCTGACGATCAGATGCGCGCATCCGCTTCCG 4447
QY 1081 GTGCGGAACGGCGAGCGCGATCAGCGCACCGGACGCGCGCGCGCGCGATCAGCATTC 1140
DB 4448 GTGCGGAACGGCGAGCGCGATCAGCGCACCGGACGCGCGCGCGCGCGATCAGCATTC 4507
QY 1141 ATGCTCAT 1149
DB 4508 ATGCTCAT 4516
RESULT 7
AR169831
LOCUS AR169831 8625 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6291204.
ACCESSION AR169831
VERSION AR169831.1 GI:17907751
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
Unclassified.
1 (bases 1 to 8625)
AUTHORS Pasamontes,L. and Teygankov,Y.
TITLE Fermentative carotenoid production
JOURNAL Patent: US 6291204-A 1 18-SEP-2001;
FEATURES
location/Qualifiers
source
1..8625
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1149; DB 6; Length 8625;
Best Local Similarity 100.0%; Pred. No. 1.8e-141;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGCTCTCTCTGACAGGAGGCGTTGCGGACAGCGGAGCGCGCTGCGACAGCGG 60
DB 5794 TCATGCTCTCTCTGACAGGAGGCGTTGCGGACAGCGGAGCGCGCTGCGACAGCGG 5853
QY 61 AATGGGCGGCGCTGCGGATGAGCATGCGAAGCGGTCGCGCAATGTCAAGCGCCCGGCATA 120
DB 5854 AATGGGCGGCGCTGCGGATGAGCATGCGAAGCGGTCGCGCAATGTCAAGCGCCCGGCATA 5913
QY 121 GAAGCGCTCGATCAGCGGCTGCGGAGCGGCTGAGAAACGCTTGCAGCAGGCGCATAGCAGC 180
DB 5914 GAAGCGCTCGATCAGCGGCTGCGGAGCGGCTGAGAAACGCTTGCAGCAGGCGCATAGCAGC 5973
QY 181 GTCGCGCGGCGAGCGCGGGAACAGCATCCGCTTACGAGCGCGGAGGAGGCGGTCGCGATC 240
DB 5974 GTCGCGCGGCGAGCGCGGGAACAGCATCCGCTTACGAGCGCGGAGGAGGCGGTCGCGATC 6033
QY 241 CGCGGATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATGAGGATTCGCG 300
DB 6034 CGCGGATGATGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGATGAGGATTCGCG 6093
QY 301 CGCGGATGAGCATCCGCGACCTGCGCGGATAGGCGAGCGAATTCGCGTGAACGCGGAG 360

Db 6094 CGCCGGATGATCCGCGCACTGCGCGCATAGGCGACGATATCCGGTGAAGGGGTG 6153
QY 361 GAACAGCCCTGCCCCAGCCCAACCGGCAACCGGCCCCCTGGCTGGTCCGCCAGAAAGCC 420
Db 6154 GAACAGCCCTGCCCCAGCCCAACCGGCAACCGGCCCCCTGGCTGGTCCGCCAGAAAGCC 6213
QY 421 TATGGCGTATGGGCGCAGGCGCATGAGGAGGAGATGCCCTTTTCGCGCGCATCTCTGCCC 480
Db 6214 TATGGCGTATGGGCGCAGGCGCATGAGGAGGAGATGCCCTTTTCGCGCGCATCTCTGCCC 6273
QY 481 GGTCCAGCCCGCTGCGCGCATATGTCACAGCAAGCTTCGCGCAAGCGCCCATGCTCCAG 540
Db 6274 GGTCCAGCCCGCTGCGCGCATATGTCACAGCAAGCTTCGCGCAAGCGCCCATGCTCCAG 6333
QY 541 ATGCGCGCGCTGCTAGCGCGTATCTTCGATCAGAGTGGCGGTGGAGCTGAAGGGCGAG 600
Db 6334 ATGCGCGCGCTGCTAGCGCGTATCTTCGATCAGAGTGGCGGTGGAGCTGAAGGGCGAG 6393
QY 601 CAGATGATGAAAGCGGTACCCGTCATCTGCGGAAAGGTTCGCGCATGATCATCGGGCG 660
Db 6394 CAGATGATGAAAGCGGTACCCGTCATCTGCGGAAAGGTTCGCGCATGATCATCGGGCG 6453
QY 661 CTGCAAGCGCATGGGCGGCGCTGCTGATCTTGACGCCCAAGATTTCTGAAACCCAC 720
Db 6454 CTGCAAGCGCATGGGCGGCGCTGCTGATCTTGACGCCCAAGATTTCTGAAACCCAC 6513
QY 721 GGTCAAGGTTCGGGTCTTCACACGCAACGCGGTGATCAGCAGGCAAGCTTCGATCCG 780
Db 6514 GGTCAAGGTTCGGGTCTTCACACGCAACGCGGTGATCAGCAGGCAAGCTTCGATCCG 6573
QY 781 CGAGCGCTCGTCAAGCGTCCGCGGTATCTGTCAGCGTCCAGCATGCGTATTCACCG 840
Db 6574 CGAGCGCTCGTCAAGCGTCCGCGGTATCTGTCAGCGTCCAGCATGCGTATTCACCG 6633
QY 841 CAGATGACACCTTGACGACAGCCGATCAGCGCGCCGCTCGATGAGCCATAGCTCT 900
Db 6634 CAGATGACACCTTGACGACAGCCGATCAGCGCGCCGCTCGATGAGCCATAGCTCT 6693
QY 901 CGTCAAGCGCGCGGATGCTGCGGAAACGCGACTCTCTGATCCGTTCATTCGCGCGAG 960
Db 6694 CGTCAAGCGCGCGGATGCTGCGGAAACGCGACTCTCTGATCCGTTCATTCGCGCGAG 6753
QY 961 AATGGGCGACAGGCGCGCGCATTCGCGGCGAAATCGGTGCTGCGGAGGACAGG 1020
Db 6754 AATGGGCGACAGGCGCGCGCATTCGCGGCGAAATCGGTGCTGCGGAGGACAGG 6813
QY 1021 GTGCTGCTCGAGGCGCGGACCGCGCTGAGCATCAGATGCGGCAATCCGCTGCG 1080
Db 6814 GTGCTGCTCGAGGCGCGGACCGCGCTGAGCATCAGATGCGGCAATCCGCTGCG 6873
QY 1081 GTTCGGAAGCGGAGCGGATCAGCGGACCGGACAGCCCGCGCGCGCATCAGCGATC 1140
Db 6874 GTTCGGAAGCGGAGCGGATCAGCGGACCGGACAGCCCGCGCGCGCATCAGCGATC 6933
QY 1141 ATGGCTCAT 1149
Db 6934 ATGGCTCAT 6942

RESULT 8
AR452005
LOCUS AR452005 8625 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1 from patent US 6677134.
ACCESSION AR452005
VERSION AR452005.1 GI:42683332
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8625)
AUTHORS Paramotes, L. and Tsygankov, Y.
TITLE Fermentative carotenoid production
JOURNAL Patent: US 6677134-A 1 13-JAN-2004;

FEATURES
source
Location/Qualifiers
1..8625
/organism="unknown"
/mol_type="genomic DNA"
Query Match 100.0%; Score 1149; DB 6; Length 8625;
Best Local Similarity 100.0%; Pred. No. 1,8e-141;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
QY 1 TCAATGCTCTCTCTGACGAGGAGGCGTTGCGGCGAGCGGACGCGCTGCGACAGCGG 60
Db 5794 TCAATGCTCTCTCTGACGAGGAGGCGTTGCGGCGAGCGGACGCGCTGCGACAGCGG 5853
QY 61 AATGGGCGGCGCTCGGCTGACGATGGAAGCGGCTCGGCAATGTCAGGCGCCGCGCAT 120
Db 5854 AATGGGCGGCGCTCGGCTGACGATGGAAGCGGCTCGGCAATGTCAGGCGCCGCGCAT 5913
QY 121 GAAGCGCTGATCAGCGGCTGCGGAGCGGTAGAACCGCTGACAGCGCATAGCGAG 180
Db 5914 GAAGCGCTGATCAGCGGCTGCGGAGCGGTAGAACCGCTGACAGCGCATAGCGAG 5973
QY 181 GTGCGGCGGCGAGCGCGGAGACATCCGTTTCAAGACGCCGAGAAACGCTCGCATC 240
Db 5974 GTGCGGCGGCGAGCGCGGAGACATCCGTTTCAAGACGCCGAGAAACGCTCGCATC 6033
QY 241 CGCGGATGATAGGCGCGCGCGGACCGCGGAGCGGCGGAGCGGCTGTCAGGTCGCG 300
Db 6034 CGCGGATGATAGGCGCGCGGACCGCGGAGCGGCGGAGCGGCTGTCAGGTCGCG 6093
QY 301 CGCGGATGATAGGCGCGGACCGGCTGCGGAGCGGATAGGAGCGGATATCCGATAGCGG 360
Db 6094 CGCGGATGATAGGCGCGGACCGGCTGCGGAGCGGATAGGAGCGGATATCCGATAGCGG 6153
QY 361 GAACAGCCCTGCGCCCAAGCCGACCGCGCCCTCTGCGGTGTCGCGCGCAAGAGCC 420
Db 6154 GAACAGCCCTGCGCCCAAGCCGACCGCGCCCTCTGCGGTGTCGCGCGCAAGAGCC 6213
QY 421 TATGGCTATAGGCGGACGAGGATGAGGATGCCCCCTTTGCGCGCGGATCTCTGCGCC 480
Db 6214 TATGGCTATAGGCGGACGAGGATGAGGATGCCCCCTTTGCGCGCGGATCTCTGCGCC 6273
QY 481 GGTCCAGCCCGCGCTGCGGCGCATATGTCAGGACGCTGCGCGAGCGCCATCGTCAG 540
Db 6274 GGTCCAGCCCGCGCTGCGGCGCATATGTCAGGACGCTGCGCGAGCGCCATCGTCAG 6333
QY 541 ATGCGCGCGCTGCTGTAGCGGCTATCTTCGATCAGATGCGGCTGAGGACTGAAGGCGAG 600
Db 6334 ATGCGCGCGCTGCTGTAGCGGCTATCTTCGATCAGATGCGGCTGAGGACTGAAGGCGAG 6393
QY 601 CAGATGATGAAAGCGGTACCCGTCATCTGCGGAAACGCTGCGGTCAATGATCGGGG 660
Db 6394 CAGATGATGAAAGCGGTACCCGTCATCTGCGGAAACGCTGCGGTCAATGATCGGGG 6453
QY 661 CTGCAAGCGCGGCGGCTCGGCTGTCGATCTCGACGCCCAAGAAATTTCTGAAACCCAC 720
Db 6454 CTGCAAGCGCGGCGGCTCGGCTGTCGATCTCGACGCCCAAGAAATTTCTGAAACCCAC 6513
QY 721 GGTCAAGGTTCGGGTCTTCACGCGCAACAGGCGGTGATCAGCGAGCGCTTCGATCCG 780
Db 6514 GGTCAAGGTTCGGGTCTTCACGCGCAACAGGCGGTGATCAGCGAGCGCTTCGATCCG 6573
QY 781 CGAGCGCTCGTCAAGCGTCCGCGGTATCTGTCAGCGTTCGGAACATGTTTCCACCG 840
Db 6574 CGAGCGCTCGTCAAGCGTCCGCGGTATCTGTCAGCGTTCGGAACATGTTTCCACCG 6633
QY 841 CAGATGACACCTTGACGACAGCCGATCAGCGCGCGCTCGATGAGCGCATAGCTCT 900
Db 6634 CAGATGACACCTTGACGACAGCCGATCAGCGCGCGCTCGATGAGCGCATAGCTCT 6693
QY 901 CGTCAAGCGCGCGGATGCTGCGGAAACGCGACTCTTGATCGGTTCATTCGCGCGAG 960
Db 6694 CGTCAAGCGCGCGGATGCTGCGGAAACGCGACTCTTGATCGGTTCATTCGCGCGAG 6753

Qy	961	AATGGGCGACAGGGCGCGCACCATTTGGGGGGGAAAGATCCGTGTGCGAGGACACAGGT	1020
Db	6754	AATGGGCGACAGGGCGCGCACCATTTGGGGGGGAAAGATCCGTGTGCGAGGACACAGGT	6813
Qy	1021	GTGCTGTGTCCAGAGGGCGCGACCGCGCGCTCGAGCATCAGATGCGCGCATCCGGTCTGCG	1080
Db	6814	GTGCTGTGTCCAGAGGGCGCGACCGCGCGCTCGAGCATCAGATGCGCGCATCCGGTCTGCG	6873
Qy	1081	GTCCGGAACGGCAGCGCGATCATCAGCGCACCCGGAACAGCCCCCGCCCGCGATCAGCAATC	1140
Db	6874	GTCCGGAACGGCAGCGCGATCATCAGCGCACCCGGAACAGCCCCCGCGCGATCAGCAATC	6933
Qy	1141	ATGGCTCAT	1149
Db	6934	ATGGCTCAT	6942
RESULT 9			
AR169852			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best local similarity			
Matches 1149; Conservative			
100.0%; Score 1149; DB 6; Length 11233;			
100.0%; Pred. No. 1.7e-141;			
Mismatches 0; Indels 0; Gaps 0;			
Qy	1	TCATGCTTCTCTCTGACAGAGGGGCGCTTCGGGAGCAGCGCGCAACGGCTTGCAACAGCG	60
Db	6476	TCATGCTTCTCTCTGACAGAGGGGCGCTTCGGGAGCAGCGCGCAACGGCTTGCAACAGCG	6535
Qy	61	AATGGGCGGGCGCTCCGGTGAAGATGCCAAGCCGGTCCGACCAATGTCAAGCGCCCGGCATA	120
Db	6536	AATGGGCGGGCGCTCCGGTGAAGATGCCAAGCCGGTCCGACCAATGTCAAGCGCCCGGCATA	6595
Qy	121	GAAGCGCTCATGTCAGCGGCGCTCGGAGCGGTTAGAACCGCTGACAGACGAGCGATAGCAGC	180
Db	6596	GAAGCGCTCATGTCAGCGGCGCTCGGAGCGGTTAGAACCGCTGACAGACGAGCGATAGCAGC	6655
Qy	181	GTCGGGCGGGCAGCCGCGGAACAGCATCCGTTACAGACCGCGCAAGAAAGCGTTCGCATC	240
Db	6656	GTCGGGCGGGCAGCCGCGGAACAGCATCCGTTACAGACCGCGCAAGAAAGCGTTCGCATC	6715
Qy	241	CGCCCGATTCATGATGCGCCAGCCCGCGCACCGCGCGAACCGGGCGGACCGCGTCTCAAGTCCG	300
Db	6716	CGCCCGATTCATGATGCGCCAGCCCGCGCACCGCGCGAACCGGGCGGACCGCGTCTCAAGTCCG	6775
Qy	301	CGCCCGATTCATTCGCGACCTTGCGCGGACATTAAGGACAGCGAATATCCGATGACGGGGTG	360
Db	6776	CGCCCGATTCATTCGCGACCTTGCGCGGACATTAAGGACAGCGAATATCCGATGACGGGGTG	6835
Qy	361	GAACAGCCCTTCCCGCAGCCCGCAACCGGACCGCCCGCTGCGGATGCGCCAGAAACC	420
Db	6836	GAACAGCCCTTCCCGCAGCCCGCAACCGGACCGCCCGCTGCGGATGCGCCAGAAACC	6895
Qy	421	TATGGCGTCATAGGCGCACCGCGATGAGGACAGATGCCCTTTTCGCGCGCATCTCTCTGCC	480
Db	6896	TATGGCGTCATAGGCGCACCGCGATGAGGACAGATGCCCTTTTCGCGCGCATCTCTCTGCC	6955

QY	481	GGTCAGAGCCCGCGCTGGGCGGATAGTGCAGGACGCCCTGGCGCAGCGCGGCATGCTCAG	540
Db	6356	GGTCAGAGCCCGCGCTGGGCGGATAGTGCAGGACGCCCTGGCGCAGCGCGGCATGCTCAG	7015
QY	541	ATCGCCGCGCTCGCTGAGCGCGTATCTCGATCAGGATGCGGAGTGGAGTGAAGGCGAG	600
Db	7016	ATCCGCCCGCGCTGAGCGCGTATCTCGATCAGGATGCGGAGTGGAGTGAAGGCGAG	7075
QY	601	CAGATTAGATGAAGGCGGTAAACCGGTCATCTGGCGGAAACGGTGGCGTCATGATCATCGGCG	660
Db	7076	CAGATTAGATGAAGGCGGTAAACCGGTCATCTGGCGGAAACGGTGGCGTCATGATCATCGGCG	7135
QY	661	CTCGACGCCCATGGGCGGCGGTGCTCGATCTCGATCTCGACGCCCAAGATTCTTGAAAACCCAC	720
Db	7136	CTCGACGCCCATGGGCGGCGGTGCTCGATCTCGATCTCGACGCCCAAGATTCTTGAAAACCCAC	7195
QY	721	GGTCAGAGTGGCGGGGCTTCGACACGGGACCAACGGGCGTCCATTCACGAGGACGCTGATCCG	780
Db	7196	GGTCAGAGTGGCGGGGCTTCGACACGGGACCAACGGGCGTCCATTCACGAGGACGCTGATCCG	7255
QY	781	CGACCCGCTCGTCAGCGCTGCGCGCGGTATGTCAGAGGTCGCGCATGTCGATTTCCACCG	840
Db	7256	CGACCCGCTCGTCAGCGCTGCGCGCGGTATGTCAGAGGTCGCGCATGTCGATTTCCACCG	7315
QY	841	CAGATTCGACACCTTCGACGACGCCGATCAGCGCGCCCGCTCGATTCGAGCCATAGCTGT	900
Db	7316	CAGATTCGACACCTTCGACGACGCCGATCAGCGCGCCCGCTCGATTCGAGCCATAGCTGT	7375
QY	901	CGTCAGAGCGGCGCGAATGTCGCGGAAACGCACTTCCTGATCCGTCATTTGCCGCGACG	960
Db	7376	CGTCAGAGCGGCGCGAATGTCGCGGAAACGCACTTCCTGATCCGTCATTTGCCGCGACG	7435
QY	961	AATGCGCGACAGCGCGCGCACCTTTCGGCGGAAAGATCCGTCGTCGGCGAGACACGTT	1020
Db	7436	AATGCGCGACAGCGCGCGCACCTTTCGGCGGAAAGATCCGTCGTCGGCGAGACACGTT	7495
QY	1021	GTCGTCGTCGCGAGGCGGCGGACCGCGCGGTGAGGATCACATGCGGCGCATCCGATTCGCG	1080
Db	7496	GTCGTCGTCGCGAGGCGGCGGACCGCGCGGTGAGGATCACATGCGGCGCATCCGATTCGCG	7555
QY	1081	GTCGCGAAGCGCAAGCGGATCAGCGCACCGGACAGCCCGCGCGCGCGATCAGCAATC	1140
Db	7556	GTCGCGAAGCGCAAGCGGATCAGCGCACCGGACAGCCCGCGCGCGCGATCAGCAATC	7615
QY	1141	ATGCGTCAT 1149	
Db	7616	ATGCGTCAT 7624	

RESULT	10			
AR452026				
LOCUS				
DEFINITION	AR452026	11233 bp	DNA	
ACCESSION	Sequence 27 from patent US 6677134.			linear
VERSION	AR452026			
KEYWORDS	AR452026.1	GI:42683353		
SOURCE	.			
ORGANISM	Unknown.			
	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 11233)			
TITLE	Paramontes, L. and Tsygankov, Y.			
JOURNAL	Fermentative carotenoid production			
FEATURES	Patent: US 6677134-A 27 13-JUN-2004;			
	Location/Qualifiers			
source	1..11233			
	/organism="unknown"			
ORIGIN	/mol_type="genomic DNA"			
Query Match	100.0%;	Score 1149;	DB 6;	Length 11233;
Best Local Similarity	100.0%;	Pred. No. 1,7e+141;		
Matches 1149;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Db 861 GGACAGACCCGCCACCACTGCCCACTGTCCGATAGGGCAGAGATAGCCGCTGAC 802
QY 355 GGGGTGGAAACAGCCCTGCCCCAGCCGCAACCGGACCGCCCTGTGGTGTGGGCA 414
Db 801 CGGATGAAAGAACCCCGCGCGAGTCCACGGGAACAGGCCCGCGCGGTGATCGGCCA 742
QY 415 GAAGCCTTAGCGCTCATAGGAGCAGCGCATAGGAGCAGAGATGCCCTTTTCGCGCCGATCTC 474
Db 741 GAAGCCCGCGCATATAGGAGCAGCGCATAGGAGATAGCGCGCTTTCGCGCCGACCTC 682
QY 475 CTGCCCCGTTCACCCCGCTTGGCCGCGCATATGTCACGCCACCGCTTGGCCGCGCATC 534
Db 681 GGGCCCGGTTCAGCCCTGTGGCGGGCATATGTCGAGGAGCGCCGCGCAGCGCGTCTC 622
QY 535 GTCCAGATGCGCCCGCTGCTGATAGCGGCTATCTTCATCAGATAGCGGGTGGAGCTGAA 594
Db 621 GTCCAGATGCGCCCATAGGAAATAGCGCTGTCTTCATCAGATAGCGGGTGGAGAGAA 562
QY 595 GGGCAGCAGATAGATAGAACCGGTACCCGTCATCTTCGGAACGCTGCGCTCATGATCAT 654
Db 561 GGGCAGCAGATAGATAGAACCGGTACCCGTCATCTTCGGAACGCTGCGCTCATGATCAT 502
QY 655 CGGGCGCTTCAGACCCCATAGGGGGGCTTCGCTTCATCTTCAGACGCCCATATTTCTGAA 714
Db 501 CGGGCGGGGACAGCGCTGGGGGGCGGTCTCTCATCTTCAGACCCCATATTTCTGAA 442
QY 715 ACCCAAGCTCAGATGCGGGGTCTTCAGCGGACCAACGGGCTGATCAGCAGCAGACCTTC 774
Db 441 ACCCAAGCTCAGATGCGGGGTCTTCAGCGGACCGCCCGCTTCAGGACCGCGCGCTTC 382
QY 775 GATCCGCGAGCCGCTCCGTCAGCGCTCGCGCGGTATCTTCAGCGCTTCGATTCGATTT 834
Db 381 GATCCGCGAGCTCCGTCAGCGCTCGCGCGGTATCTTCAGCGCTTCGATTCGATTT 322
QY 835 CCAACCGCATATGACACCTCTGACAGAC-----CGATCAGCGCGCCCGCTTCATGCA 888
Db 321 CCAAGCGATCTCGGCGCGCTCGACCGACCGCATCTCGCGCGCGCGCTTCAGCGCA 262
QY 889 GCCATAGCTCTGCTCAGCGCGCGCGCATATGTCGAGGAAACGCGACCTCTTCATCTCTCA 948
Db 261 CCGCTTACCGGTGTGCGCGCGCGCGCATATGCGGGGAAAGCGACCTCTTCATCTCTCA 202
QY 949 TTCGCGCGCAGCATATGAGGCGACAGCGCGCGCGCATATTCGAGCGGAAATCTCTGTCTG 1008
Db 201 GTTGGCGCGCGAGAGGCGCTTCAGCGCGCGCGCATCTCGGCGGACAGAGTCTGGGTCG 142
QY 1009 GCAAGACCAAGTGTCTGTGTCGAGGGGGCGGACCGCGCTGAGCATATCAGATCGCGCG 1068
Db 141 GCAAGACCAAGTGTGTGCGCTGTGACCGGTCTTCGCGCATATGTCAGGACGACACCGCG 82
QY 1069 ATCCGGTCTGCGGTGCGGAAAGCGGACGCGATCAGCGGACGCGCGCGCGCGCGCG 1128
Db 81 GTCCGGCTC 22
QY 1129 GATCAGCAGATCATGGCTCAT 1149
Db 21 CAGCAGCAGCTCATGGCTCAT 1

RESULT 12
AR041163/c 1161 bp DNA linear PAT 29-SEP-1999
LOCUS AR041163
DEFINITION Sequence 5 from patent US 5811273.
ACCESSION AR041163
VERSION AR041163.1 GI:5961659
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1161)
AUTHORS Misawa,N., Kondo,K., Kajiwara,S. and Yokoyama,A.
TITLE DNA strands useful for the synthesis of xanthophylls and the process for producing the xanthophylls

JOURNAL Patent: US 5811273-A 5 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..1161
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 55.8%; Score 640.8; DB 6; Length 1161;
Best Local Similarity 73.4%; Pred. No. 1e-74;
Matches 851; Conservative 0; Mismatches 297; Indels 12; Gaps 2;
QY 1 TCATCTCTTCTCTCTCAGACAGGGGCGCTTCGGGCGAGGCGACAGCGCTTCGACAGCGG 60
Db 1161 TCATCTCTTCTCTCTCAGACAGGGGCGCTTCGGGCGAGGCGAGCGAGCGCTTCGACAGGGG 1102
QY 61 AATGGCGGGGCTGCGGAGCATGCGAAAGCGGCTGCGCAATGTCAGAGCGCGCGCATTA 120
Db 1101 AATGGAGAGCTTTCGCGGTCAGATGCGAGCTGATTCGCGCAGCTCAGCGCGCGCATTA 1042
QY 121 GAAGCGCTTCATCAGCGGCTGCGGAGCGGTAGAACCGCTTCAGCAGGCGCATAGCGAG 180
Db 1041 GAACGCTTCATCAGCTCATGCGGATGCGGAGAACCGCTTCAGCAGGATATAGCGCG 982
QY 181 GTCCGGCGGGGACCGCGGGAACGATTCGGTTACGACCGCGCAGAGAGCGGTGCGCATC 240
Db 981 GTCCGGCGCGGACCGCGGGAACGATTCGGTTACAAAGGCGCAGAAAGCGGTGCGCGG 922
QY 241 CGCGGATCGATGCG 294
Db 921 CGCGGATCGATGCG 862
QY 295 GTCCGCGCGCGCGGATGCGCATCTGCGCGCGCATAGGGCAGCGAATATCCGCTGAC 354
Db 861 GGAACAGACCCCGCACCATGTCGCGCACCTGTGCGCGCATAGGGCAGCGAATAGCGGTGAC 802
QY 355 GGGGTGAAACAGCCCTGCCCCAGCCGCAACCGGACCGCGCTTCGCGGTGTGCGCGCA 414
Db 801 CGGATGAAAGAACCCCGCGCGAGTCCACGGGAAACAGGCCCGCGCGGTATTCGCGCA 742
QY 415 GAAGCCTTAGCGCTCATAGGAGCAGCGCATATGAGGAGGAGCGCCCTTTTCGCGCGCATCTC 474
Db 741 GAAGCCCGCGCATATAGGAGCAGCGCATATGAGGAGGAGATGCGCGCTTCGCGCGCACTC 682
QY 475 CTGCGCGGTTCAGACCCCGCTTCGCGCGCATATGTCAGCGACCGCTTCGCGCGCGCATC 534
Db 681 GGGCCCGGTTCAGACCCCGCTTCGCGCGCATATGTCAGCGACCGCGCGCGCGCGCGCTC 622
QY 535 GTCCAGATGCGCGCGCTGTGTCGAGCGGTATCTTCATCAGATAGCGGGTGGAGCTGAA 594
Db 621 GTCCAGATGCGCGCGCATGGAATAGCGGTGTCTTCATCAGATAGCGGGTGGAGGAA 562
QY 595 GGGCAGCAGATAGATAGAACCGGTACCCGTCATCTTCGGAACGCGTGCCTCATGATCAT 654
Db 561 GGGCAGCAGATAGATAGAACCGGTACCCGTCATCTTCGGAACGCGTGCCTCATGATCAT 502
QY 655 CGGGCGCTTCAGACCCCATAGGGGGCGTTCGCTTCATCTTCAGACGCCCATATTTCTGAA 714
Db 501 CGGGCGGGGACAGCGCTGGGGGGCGGTCTTCATCTTCAGACCCCATATTTCTGAA 442
QY 715 ACCCAAGCTCAGATGCGGGGTCTTCAGCGGACCAACGGGCGCTTCAGCAGGAGCGCTC 774
Db 441 ACCCAAGCTCAGATGCGGGGTCTTCAGCGGACCGCGCGCGCGCGCGCGCGCGCGCTC 382
QY 775 GATCCGCGAGCGCTCCGTCAGCGCTGCGCGGTATCTTCAGCGCTGCGACATGCGTATT 834
Db 381 GATCCGCGGTCCGCGAGGACAGGCTGCGCGCTTCGCGCATTCAGAGGAGGATGCTGCTGTC 322
QY 835 CCAACCGCATATGACACCTTCGACAGC-----CGATCAGCGCGCGCGCGCTTCATGCA 888
Db 321 CCAAGCGATCTCGGCGCGCGCGCGCGCGCATTCGCGACGCGCGCGCGCGCGCTTCAGGCA 262
QY 889 GCCATAGCTCTGCTCAGCGCGCGCGCATATGTCGAGGAAACGCGACCTCTTCATCTCTGCA 948

Db 261 CCCGTAACCGGTGCGCAGCCCGGCGATGCGGGGAAAGCGACCTCTGTGCGGCGCA 202
QY 949 TTGCGCGCGACGAAATGGGCGGAGCGCGCGACGCACTTGGGGGAAAGATCCGTGTCTG 1008
Db 201 GTTGGCGCGCGCGAGGCGGCTTTCAGCCGCGCGACGCGATCCGGGACAGGTGGGGTCTG 142
QY 1009 GCAGACACGAGTGTGCTGTGCTCCGAGGGGCGCGACCGCGCTCGAGATCAAGATGCGCG 1068
Db 141 GCAGACACGAGTGTGCGCGCTGTGACGCTCTGCGCGCATGCTCCAGCAGACGCGCGAG 82
QY 1069 ATCCGGTCTGCGGTCCGGAACGCGCAACCGGATCAGCGACCGGACAGCCCGCGCGCG 1128
Db 81 GTCCG 22
QY 1129 GATCAGACAGATCATGCGCTCA 1148
Db 21 CAGCAGACGATCATGCGCTCA 2

RESULT 13
AR082017/c 1161 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 5 from patent US 5972690.
ACCESSION AR082017
VERSION AR082017.1 GI:10008743
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1161)
AUTHORS Misawa,N., Kondo,K., Kajiwara,S. and Yokoyama,A.
TITLE DNA strands useful for the synthesis of xanthophylls and the process for producing the xanthophylls
JOURNAL Patent: US 5972690-A 5 26-OCT-1999;
FEATURES
source 1.1161.
Location/Qualifiers
1..1161
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 55.8%; Score 640.8; DB 6; Length 1161;
Best Local Similarity 73.4%; Pred. No. 18-74;
Matches 851; Conservative 0; Mismatches 297; Indels 12; Gaps 2;
QY 1 TCATGCTCTCTCTCTGACGAGGGGGGCTTGGGCGAGCGACCGCGCTGCGACAGCGG 60
Db 1161 TCATGCTCTCTCTCTGACGAGGGGGGCTTGGGCGAGCGACCGCGCTGCGACAGCGG 1102
QY 61 AATGGGCGGGCGGTCCGGTGCAGATGCGAAGCGGTGCGGCAATGTCAAGGCGCGCGCAT 120
Db 1101 AATGGGCGGGCGGTCCGGTGCAGATGCGAAGCGGTGCGGCAATGTCAAGGCGCGCGCAT 1042
QY 121 GAAGCGCTGCATCAGCGGCTGCGGCGAGCGGTAGAAACGCTGCGAGCGAGCGATAGCGAG 180
Db 1041 GAACGCTGCATCAGTCATGCGGCGATGCGGTAGAAACGCTGCGAGCGAGGTATAGCGCG 982
QY 181 GTCCG 240
Db 981 GTCCG 922
QY 241 CGCGCGATGCATGAGCG 294
Db 921 CGCGCGATGCATGAGCG 862
QY 295 GTCCG 354
Db 861 GGAACAGACCG 802
QY 355 GGGGTGGAACAGCG 414
Db 801 CGGATGAAGAAACCG 742
QY 415 GAAGCTTAGGCGCTCATGCG 474

Db 741 GAAGCCCGCGCGATCATGGGCGAGCGCGATGGGAAAGATGCCGCTTCCGCGCGGACCTC 682
QY 475 CTGCGCGCTTCAGCCCG 534
Db 681 GCGCCCGGTTCAGCCCG 622
QY 535 GTCCAGATGCG 594
Db 621 GTCCAGATGCG 562
QY 595 GGGCAGCAGATGATGAAGCGGTACCGGTCATCTTCGCGAAGCGTGCCTCATGATCAT 654
Db 561 GGGCAGCAGATGATGAAGCGGTACCGGTCATCTTCGCGAAGCGTGCCTCATGATCAT 502
QY 655 CGGCGCTGCAGCG 714
Db 501 CGGCG 442
QY 715 ACCCAGCGTCAGGTGCGGGGTCTCGACCGGACACAGGCGGTGCATCAGCAGGCGACCTC 774
Db 441 ACCCAGCGTCAGGTGCGGGGTCTCGACCGGACACAGGCGGTGCATCAGCAGGCGACCTC 382
QY 775 GATCCGCGAGCGCGTCCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 834
Db 381 GATCCGCGAGCG 322
QY 835 CCAAGCGCAGATGACACCGCTGCGACAGC-----CGATCAGCGCGCGCGCGCGCGCG 888
Db 321 CCAAGCGCAGATGACACCGCTGCGACAGC-----CGATCAGCGCGCGCGCGCGCGCG 262
QY 889 GCGATGCTGCTGCTGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948
Db 261 CCGTACCGGTGCG 202
QY 949 TTCCGCGCGACGAAATGGGCGGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1008
Db 201 GTTGGCG 142
QY 1009 GCAGACACGAGTGTGCTGTGCTCCGAGGGGCGCGACCGCGCTGCGAGATCAAGATGCGCG 1068
Db 141 GCAGACACGAGTGTGCTGTGCTCCGTGACCGCTCTCGCGCATGCTCCAGCAGACGCGCGAG 82
QY 1069 ATCCGGTCTGCGGTCCGGAACGCGCAACCGGATCAGCGACCGGCGCGCGCGCGCG 1128
Db 81 GTCCG 22
QY 1129 GATCAGACAGATCATGCGCTCA 1148
Db 21 CAGCAGACGATCATGCGCTCA 2

RESULT 14
E22378 1161 bp DNA linear PAT 18-JUN-2001
LOCUS E22378
DEFINITION beta-Carotene hydroxylase gene and use thereof.
ACCESSION E22378
VERSION E22378.1 GI:13024037
KEYWORDS JP 1999046770-A/2.
SOURCE JP 1999046770-A/2.
ORGANISM unidentified
unclassified
unclassified.
REFERENCE 1 (bases 1 to 1161)
AUTHORS Norihiro,M., Kazumori,M., Kiuchi,K. and Hiroyuki,F.
TITLE beta-Carotene hydroxylase gene and use thereof
JOURNAL Patent: JP 1999046770-A 2 23-FEB-1999;
KIRIN BREWERY CO LTD
OS Agrobacterium aurantiacum
PN JP 1999046770-A/2
PD 23-FEB-1999
PF 07-AUG-1997 JP 1997213648
PR NORIHIKO MISAWA,KAZUMORI MASAMOTO,KIUCHI KANEKO,HIROYUKI FUJI

[illegible]

QY	775	GATCCGCGCAGCCGCTCCGTCACGCTGCGCGCGCGATATCGTCCAGCGCTCGGACATCGTAAT	834
Db	381	GATCCGGGTGCCGACAGGACACGCGTGTGCCCCCTGGCGCATCCAGAGGGGAGATGTCGCTCTC	322
QY	835	CCACCGCAGATCGACACCTTCGACGACG-----CCGATCAGCGCGCCCGCTTCGATCGA	888
Db	321	CCAGCGGATCTCGCGCGCCCGACCGGACCAACCGCATTCGCGCCAGCGCCCGCCGTCAGCGA	262
QY	889	GCCATAGCGTGTGTGACAGGCGGCGCGGAAATGTCGGGAAACCGCGACCTTCGATCCGTCGA	948
Db	261	CCCGTAACCGGTGTGCGACGCGCGCGCGCATGGCGGAGAAAGCGCACCTTCGTGTGGGCGCA	202
QY	949	TTCCGCGCGCAGAAATGGGCGCAGACGCGCGCCAGCGCATTCGCGCGCAGAAAGATCCGTGCTG	1008
Db	201	GTTGGCGCGCGCGCAGGCGGCTTCACGCGCGCCAGCCAGTCGCGCGACAGTGGGGGTCTGT	142
QY	1009	GCAGAGCCAGGATGTGCTGTGTCGACGAGGCGCGCAGACCGCGCGTGTGAGCATCACGATCGCGC	1068
Db	141	GCAGGACCAAGATGTGCGCGCTGTGACGCGCTTCGCGGATGTGTCAGACGACGACGCGCAG	82
QY	1069	ATCCGATGTGCGGATTCGCGAACGCGAACGCGCATTCAGCGCACCGGACACGCCCGCGCCCGC	1128
Db	81	GTCGCGCGCGCGCGCGCGCGCAGCGCGCATCAGCCCGTGTGGCAAGCGCCGCGCCCTGC	22
QY	1129	GATCAGCAGATCATGCTCA	1148
Db	21	CAGCAGCAGCTCATGCTCA	2
RESULT 15			
AR041164/c	AR041164	2886 bp	DNA
LOCUS	Sequence 7 from patent US 5811273.	Linear	PAT 29-SEP-1999
DEFINITION	AR041164		
ACCESSION	AR041164.1	GI:5961660	
VERSION			
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2886)		
AUTHORS	Misawa,N., Kondo,K., Kajiwara,S. and Yokoyama,A.		
TITLE	DNA strands useful for the synthesis of xanthophylls and the process for producing the xanthophylls		
JOURNAL	Patent: US 5811273-A 7 22-SEP-1998;		
FEATURES	Location/Qualifiers		
source	1..2886		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match	55.8%;	Score 640.8;	DB 6;
Best Local Similarity	73.4%;	Pred. No. 7.4e-75;	
Matches	851;	Conservative 0;	Mismatches 297;
		Indels 12;	Gaps 2;
QY	1	TCATGCTCTCTCCCTGACAGAGGGGGCGTTGCGGAGACGGCGACCGCTTCGACACACGG	60
Db	2509	TCATGCGGTTTCTTCCTTCACAGAGGAGACGTTTCGGGAGGACGATGGCGCTGCCAAGGGG	2456
QY	61	AATGGGCGGGCGGTCCGCGTACGATTCGCGAAGCGGCTCGGCCAATCTTCAGCGCCCGGATA	120
Db	2449	AATGGAGAGCTTGGCCGGTCGACGATGCGCAGCTGATTCGCGACGCTCAGCGCGCGGATTA	2390
QY	121	GAAGCGCTCGATCAGCGCGGCTGGCGCAGCGGCTAGAACCGCTTGACAGCAGGCGATAGCAGC	180
Db	2389	GAACCGTTGCATCAGTCAGTCATCGATCGGCGCATGCGGTAGAACCGCTTGACAGCAGGCGATAGCAGC	2330
QY	181	GTCGGGCGGGGACCGCGGAAACAGCATTCGGTTACAGACCGCGCAGGAAGCGGTGGCATC	240
Db	2329	GTCGGGCGCGCGACGCGCGGAAACAGCATTCGGTTAAAGGCGGAAAGCGGTGCGCGC	2270
QY	241	CGCGCGATTCATGGCCACGCGCGCACCGCGCGACG-----GGCGGACGCGGTGTCAG	294
Db	2269	CGCCCGGTGATCGCGTAAATGCGGAAATGCGCGCGCGCGACCGCGCTTCGATGCTCGGCGCGCC	2210

This page Blank (uspto)